5-7-2016

Prediction of Internal Body Temperature using Machine Learning Models

Luke N. Belval
University of Connecticut - Storrs, luke.belval@uconn.edu

Prediction of Internal Body Temperature using Machine Learning Models

Luke N. Belval

B.S., University of Connecticut, 2014

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

At the

University of Connecticut

2016

Approval page

Masters of Science Thesis

Prediction of Internal Body Temperature using Machine Learning Models

Presented by:

Luke N. Belval, B.S.

Major Advisor____

Douglas J. Casa, Ph.D., ATC

Associate Advisor____

Lawrence E. Armstrong, Ph.D.

Associate Advisor____

Rebecca L. Stearns, Ph.D., ATC

University of Connecticut

2016

Acknowledgements

Lesley:

Your mentorship and guidance through the data collection and writing process has been invaluable. Nothing speaks louder to this than your willingness to always listen to my crazy ideas. I appreciate your innate ability to always provide a positive spin and will always value your mentorship.

Dr. Armstrong:

Your wisdom and knowledge within the field of exercise science is inspiring. I am incredibly humbled to have had you serve on my committee and have learnt so much from your classes and mentorship. I truly appreciate and value that I was able to work with you during my time at UConn.

Dr. Stearns:

You have always been so willing to help me from day one. Everything you have done behind the scene for students like me is one of KSI’s largest assets. No amount of technical support can make up for the mentorship and support you have provided me throughout the years.

Dr. Casa:

Your passion and willingness to help is contagious. I am deeply indebted to your willingness to help out a high school student and allowing for me to continue my education with such a wonderful advisor. Whenever I am asked why I would stay at UConn, it is easy to identify the reason, you.

Table of Contents

Abstract

Context: The only validated methods for assessment of deep body temperature during exercise in the heat are invasive or logistically difficult to implement. Non-invasive prediction of deep body temperature has the potential to provide critical information to individuals who exercise in environmental extremes. **Objective:** To examine the use of machine learning methods for the prediction of deep body temperature using non-invasive measures. **Setting:** Research laboratory. **Participants:** Twenty-five recreationally active participants (mean±SD; male, n=19; female, n=6, age, 24±4 y; height, 177±10 cm; body mass, 75.94±12.45 kg; body fat, 15.31±6.55%). **Interventions:** We pooled data from two studies wherein participants walked and ran on a motorized treadmill in an environmental chamber (ambient temperature, 39.8±1.7°C; relative humidity, 33.4±10.7%). 7-site skin temperature (chest, abdomen, back, upper arm, neck, thigh and calf), heart rate, speed, incline and rectal temperature were collected regularly. **Main Outcome Measures:** Data were split into a 70%/30% partition for the purposes of model development and evaluation. Skin temperature, heart rate, speed, incline, environmental conditions and demographic information were selected as predictors. Multivariate linear regression, recursive partitioning, M5' modeling and multivariate adaptive regression splines analyses were performed to develop prediction models. K-nearest neighbor and C5.0 model tree analyses were performed to develop classification models for individuals becoming hyperthermic (>39°C). **Results:** Standard stepwise linear regression accounted for 61% of the variability in rectal temperature (SEE=0.52). A Multivariate adaptive regression spline model accounted for 77.6% of the variance in rectal temperature (RMSE=0.428). A C5.0 decision tree was able to identify cases where an individual was hyperthermic with a sensitivity of 0.625 and a specificity of 0.906. This yielded a positive likelihood ratio of 6.58. **Conclusions:** Machine learning

techniques improved upon traditional regression analyses for the prediction of rectal temperature. Additionally, decision tree models were able to identify individuals who were hyperthermic with moderate shift in diagnostic probability. These techniques may be useful for refinement and implementation of future models to predict deep body temperature in an athletic setting.

I. Review of the Literature

Humans are capable of thriving in a wide variety of environmental situations. The human thermoregulatory system is theorized to have been an early advantage for humans through their practice of persistence hunting.[1] The ability to use sweat as method to dissipate the tremendous amount of heat generated by exercising muscles has allowed humans to exercise in extreme environmental conditions relatively safely.[2] In the past this exercise was necessary to eat and survive, while modernity has turned human exercise into largely an athletic or occupational pursuit. As athletes, laborers and soldiers have been pushed further the upper limits of the thermoregulatory system have been found. When an individual exceed thermoregulatory capacity, their body temperature will begin to rise. Prolonged exposure to elevated body temperature can not only hinder exercise performance, but also compromise health.[3,4]

Heat exhaustion and exertional heat stroke are the two clinical conditions that are most concerning for individuals exercising in warm environments.[5] Heat exhaustion, or the inability to continue exercise in the heat due to cardiac insufficiency, is the result of completive demand for blood flow between the exercising muscles and the skin surface as body temperature rises.[4,6] For many, this means a failure to complete a practice, game, job or mission that could have been handled in more temperate conditions. Meanwhile, exertional heat stroke is a life-threatening emergency wherein elevated body temperature leads to end-organ dysfunction.[4,7,8] There are many examples of individuals who have died during exercise as a result of exertional heat stroke.[9,10]

Generally speaking, there are three situations where exertional heat illnesses are most common: 1) American football,[11] 2) running road races, [12] and 3) military training.[13] In these situations, medical staff must be vigilant as exertional heat illnesses may present without prodromal symptoms and treatment is time sensitive. [14,15] Therefore, the prompt detection of a suspected heat illness can be considered paramount to an individual's survival. [8,16] The diagnostic criteria for exertional heat stroke are a deep body temperature greater than 40.5°C and end-organ dysfunction, typically central nervous system disturbances.[4]

Central nervous system dysfunction in itself can be readily identified, however many other conditions (e.g. hyponatremia, traumatic brain injury) cannot be ruled-out on this symptom alone.[17] This leaves deep body temperature as a critical diagnostic outcome in clinical decision making. The specific obstacles to deep body temperature assessment will be discussed further in this review, however, it is largely the invasive nature of assessment that causes clinicians hesitation.[18]

In addition to the aforementioned safety issues to performing exercise in the heat, general exercise performance has been shown to be hindered in warm environmental conditions. In fact, blinding and deceiving participants to environmental conditions has been shown to alter performance through sensory pathways.[19] Furthermore, exercise in the heat in combination of with hypohydration can cause cardiovascular drift, which in turn can hamper performance.[20,21]

Therefore, body temperature measurement could be paramount to detecting individuals who risk their safety or performance by continuing exercise while in a state of hyperthermia. In this review, the physiological mechanisms for body temperature changes will be reviewed followed by an overview of modern methods for body temperature assessment.

Exercise and Heat Stress

The human thermoregulatory system contains a homeostatic mechanism that utilizes multiple systems within the body to effectively maintain body temperature. The pre-optic area of the anterior hypothalamus contains neurons that integrate information from somatic, skin and brain temperature sensors to regulate body temperature through a system of feedback mechanisms, including eccrine sweating and vasodilation.[2] The fundamental balance of this system has been modeled through the heat balance equation: [22,23]

$$S = (M - W) \pm E \pm R \pm C \pm K$$

S = body heat storage, M = metabolic heat production, W = external work, C = convection, K = conduction, R = radiation and E = evaporation

Exertional heat stress can be labeled compensable or uncompensable heat stress. Compensable heat stress represents a heat load wherein the body can achieve a relative thermal steady state, whereas in uncompensable stress the body cannot. However, even in uncompensable heat stress, the thermoregulatory system still responds in a metered fashion in the absence of some metabolic pathology.[24,25] For example, cases of heat stroke have been attributed to exercise beyond capabilities, or situations wherein the metabolic heat production exceeds the body's ability to dissipate this heat.[10] Therefore even in the most challenging situations the body's response still can be modeled by the heat balance equation.

Through understanding the mechanisms by which the body exchanges heat, we can identify the factors that influence the rise or fall of temperatures within the body. The individual contribution of each component to heat balance is a constant dynamic during exercise that is influenced by factors both within and surrounding the body. **Figure 1** depicts the largest factors that influence the major component of the heat balance equation.[26] (Of note, Work is excluded due to the inherent link with metabolic heat production that limits its independent influence on heat balance.)

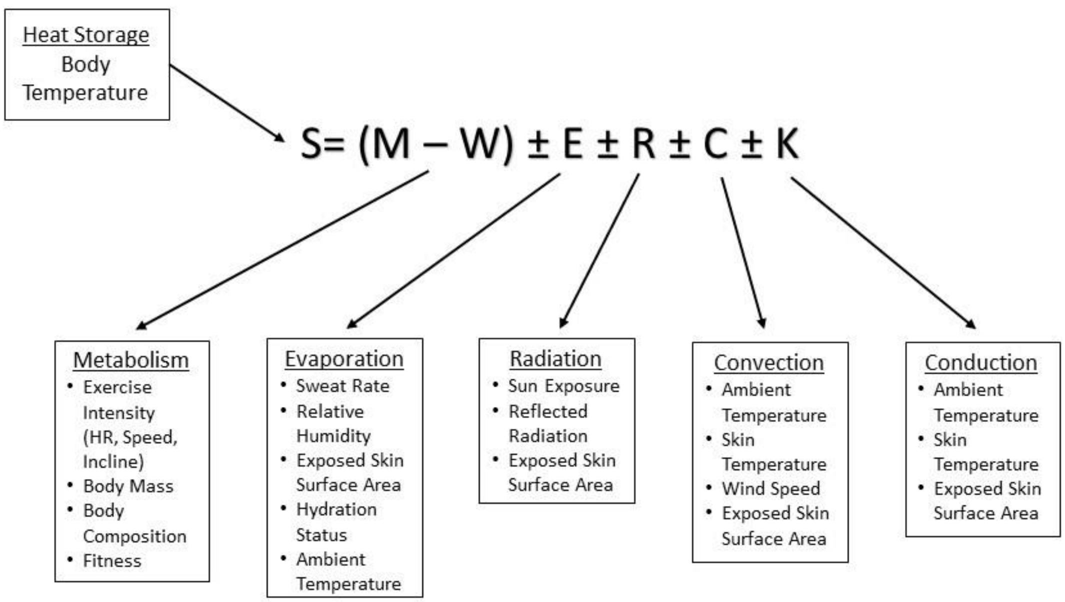

Figure 1. Factors affecting heat balance. (Adapted from Cheung, 2010)

Heat Gain

The primary mechanism for heat gain during exercise is through muscular heat production. In fact 80% of the ATP utilized by muscles is converted to heat with the remaining 20% being converted to active work.[5] Therefore the body must dissipate massive amounts of heat during even moderate exercise. The extent of this heat production is dependent on the intensity of exercise.[23] The intensity of exercise relative to heat production can be measured in several methods. The primary physiology measurement is through indirect calorimetry or VO_2. The relative VO_2 an individual consumes is directly related to exercise intensity.[27] Therefore, fitter individuals who are able to use less oxygen at a given workload may also produce less heat and sustain exercise within limits of thermal tolerance longer.

Both body mass and body composition influence the rate of heat production in several ways. First, a larger individual has to utilize a greater amount of energy for locomotion, independent of body composition.[28] Meanwhile, body composition can contribute to heat gain in two ways: 1) A high body fat percentage and 2) large lean muscle mass.[29,30] A high body fat percentage can disturb heat loss due to an alteration in the ratio of body mass to body surface area,[29] whereas a large lean muscle mass increases the total metabolic heat production.[8] It is for these reasons that in American football, the linemen are typically the individuals that have the greatest risk for exertional heat illnesses.[31]

Radiation is the only component part of the heat balance equation that contributes to heat gain consistently.[23,26] Electromagnetic radiation emitted from the sun, has a great potential to add heat to the body. Climatological strain indices, such as Wet Bulb Globe Temperature, incorporate a radiative measure to capture this effect,[32] as it can greatly alter both the perception and true measures of heat strain.[33]

Heat Loss

As mentioned previously, the evaporation of sweat is one of the most profound and unique mechanisms that humans are able to utilize. In all but the hottest and humid environments, the energy released through the vaporization of sweat is the predominant method of heat loss during exercise.[22] However, when both the thermal gradient and partial pressure of water (effectively relative humidity) gradients between the skin and the environment are diminished, i.e. hot and humid environments, the cooling power of sweat is all but eliminated.[34] These situations represent some of the highest risks for exertional heat stroke.[35] Of note, since sweat is excreted from the interstitial fluid, prolonged exercise can have great impact on hydration status which further impedes heat loss and transfer within the body.[36]

The final two components of the heat balance equation, convection and conduction, operate similarly. Both are direct dry heat transfer mechanisms, with convection being through the interaction of a fluid, commonly air. Acting primarily as heat loss mechanisms, both components can contribute to heat gain if the air temperature or clothing/equipment temperature exceeds that of the skin. Convection in particular is largely influenced by the movement of air around the individual, as the direct interfacing environment can approach a thermal equilibrium.[26] Both convection and conduction also rely on exposed skin surface area for heat transfer, with protective equipment and clothing having the potential to create microclimates that impede heat loss.[37]

The physiology of human heat balance is a dynamic process that utilizes the above stated factors to interface with the natural environment. Even through the process of heat acclimatization the increased capacity for heat dissipation is accounted for with in this basic physical understanding.[2] Therefore, our understanding of the body’s response to a given environmental and exercise stress should rely on an ability to relate the variables we study back to the basics of thermal physiology.

Body Temperature Assessment

From a clinical and performance perspective the primary variable of concern for understanding an individual's thermoregulatory state is deep body temperature. Currently the two widely accepted reference standard measures of deep body temperature are rectal and esophageal temperatures.[38,39] However, both of these measure can be considered invasive, which makes them difficult to apply in most settings for routine temperature monitoring, with many clinicians not utilizing appropriate temperature assessment.[40]

Oral, aural, forehead tympanic and axillary temperatures have been proposed as alternative methods for the measurement of body temperature, however, no study has been able to demonstrate acceptable agreement with a reference standard measure during exercise. [38,41–43] Furthermore, there is a non-uniformity in the bias for these measures, which precludes the application of a correction factor. Therefore, these non-invasive measures of body temperature are inappropriate for use in exercising individuals.

Gastrointestinal temperature, in the form of an ingestible thermistor is the only method outside of rectal and esophageal temperatures to be validated for exercising individuals. [38,42,44] However, there are several factors that impede this technique from being easily applied in the field. First and foremost, the thermistor must be in the gastrointestinal tract to signal an appropriate temperature that is not influenced by food or fluid ingestion.[38] For most individuals this means ingesting the thermistor 8-12 hours prior to the event. For diagnosis of exertional heat illnesses this is impractical. Secondly, the sensor can only be used once, creating a cost burden for clinicians and athletes trying to use the thermistor for monitoring.

Modeling of Body Temperature from Non-Invasive Measures

Researchers have attempted to non-invasively predict internal body temperature to overcome the aforementioned obstacles since 1972. Givoni and Goldman utilized standard estimates for metabolic heat production in combination with models of dry and evaporative heat loss to predict rectal temperature.[24] However Moran et al. demonstrated that this model overestimated responses in outdoor exercise situations.[45] More recent models have utilized primarily regression modeling techniques to predict body temperature based off a variety of non-invasive measures. **Table 1** presents an overview of the most relevant studies.

The occupational and military settings have been the primary outlet for the study of body temperature modeling, with only one study being performed in an athletic context.[46] To further this point, a majority of this research has been performed in equipment-laden individuals performing low to moderate intensity exercise.[47–49] The physiological responses are known to be different between a minimally clothed individual and individuals wearing equipment,[37,50] illustrating a need for the evaluation of prediction models in the former.

Table 1. Comparison of models to predict body temperature during exercise in the heat.

Study	Environmental Conditions	Exercise Protocol	Variables	Modeling Technique	Diagnostic Outcome
Xu et al., 2013	25°C, 50%RH; 35°C, 70%RH; 42°C, 25%RH; Army combat uniform with body armor; Laboratory	2h treadmill walking at 350W and 540W	Sternum T_{sk}, Sternum Heat Flux	Linear Regression	R^2=0.75
Niedermann et al., 2013	10°C, 30°C; Laboratory	Treadmill running 40% and 60% VO_{2peak}	HR, Chest Heat Flux, Back Heat Flux, Upper Arm T_{sk}, Lower Arm T_{sk}, Thigh T_{sk}	Principle Component and Linear Regression	RMSE=0.28-0.34°C
Buller et al., 2013	24-35°C, 42-97%RH; Army combat uniform with body armor; Outdoors	24h military field exercise	Heart Rate	Kalman Filter	Bias= -0.003±0.32, RMSE= 0.30±0.13
Kim et al., 2015	29.5 to 25.5°C; firefighter PPE; Laboratory	60 minutes of treadmill walking	Chest and Forehead T_{sk}	Linear Regression	T_{chest}, R^2= 0.826; $T_{forehead}$, R^2= 0.824
Richmond et al., 2015	25°C, 50%RH; 35°C, 35%RH; 40°C, 25%RH; variety of clothing conditions; Laboratory	40 minutes of walking with 20 minutes of rest	Insulated 11-site T_{sk}, microclimate T_{sk}, HR, and work	Bootstrap Regression	R^2= 0.86, SEE=0.27°C

These previously researched models utilize similar constructs for predictor selection; predictors should represent physiological processes and be relatively easily assessed. Interestingly, all but one model utilizes skin temperature. This is a likely promising candidate for a useful predictor, due to its response being tied to both the external environment and internal physiology.[51] However, the use of 11 site skin temperature sensors in conjunction with microclimate temperature as Richmond et al.,[47] may create a logistical burden that limits field application. In addition, as suggested by Buller et al., heart rate is also promising due to its interconnection with skin blood flow and hydration.[49]

Conclusion

In conclusion, body temperature is a highly important variable for individuals working, exercising or operating in the heat, whether it is to optimize performance or safety. Direct body temperature assessment currently relies on invasive or logistically difficult methods, limiting the practical application. Recent research using modeling techniques for the prediction of deep body temperature have shown promise, however, additional investigations are necessary to examine athletic populations and intense exercise with physiologically rational and logistically practical variables.

II. Introduction

Exercise in the heat is one of the most stressful situations that humans encounter. The combined environmental and exertional stressors can overwhelm the body's regulatory mechanisms and challenge both optimal performance and safety.[52] One of the greatest challenges for an athlete, laborer or soldier in this situation is combating the rise in body temperature during prolonged or intense exercise in the heat.[3,50] In particular, prolonged exposure to deep body temperatures greater than 40.5°C can result in the life threatening pathology, exertional heat stroke.[4]

Despite the known risk, in many situations actual measurement of deep body temperate is not performed, primarily due to the invasive nature of the validated measures for exercising individuals.[40] In an effort to help overcome these obstacles, researchers have attempted to utilize surface measurements as adjuncts for deep body temperature. However, to date no external device has been shown to meet acceptable limits of agreement with the reference standard measurements. [38,41–43,53]

As an alternative to direct measurement, some researchers have attempted to use mathematical models to estimate internal body temperature based upon a variety of non-invasive measures. In fact, the first models of this manner date back to 1972, where researchers used estimates of metabolic heat production, climatic conditions and clothing to predict rectal temperature.[24] More recent approaches have focused on the utilization of two primary variables, skin temperature and heart rate.[46–48,54] Of note, Buller et al. utilized a Kalman filter to predict internal body temperature from sequential heart rate measures in military exercises with an overall root mean square error (RMSE) of 0.30±0.13°C.[55] Most recently, Richmond et al. utilized

11-site skin temperature, micro-climate temperature, heart rate and work to create a prediction equation with an R^2 of 0.86 and a standard error of the estimate of 0.27°C, however, participants in their study completed a walking protocol.[56]

Machine learning is a field of computer science that develops models based on previous situations the model “learns” rather than explicit programming. In this way machine learning models are able to employ fairly simple techniques to predict more complex situations. These models can be used to either directly predict values via regression, or classify cases into subcategories.

While previous researchers have been able to obtain promising results utilizing non-invasive variables, no research has been performed on a more athletic oriented situation with intense exercise in the heat. In addition, models are specific to the situations in which they are developed and validated. Therefore, the purpose of this study was to examine different models for predicting internal body temperature based off easily accessible non-invasive measures in individuals performing strenuous exercise in the heat. Specifically, we sought to identify both predictive models and models capable of classifying individuals as hyperthermic (>39°C). It was hypothesized that using physiologically rational predictors in combination with machine learning modeling techniques could yield more useful models to predict body temperature.

III. Methods

We included data from two studies (A&B) conducted in the University of Connecticut Human Performance Laboratory. Two studies were used to provide an adequate number of data points for analysis and increase external validity. Study procedures took place in an environmental chamber (Model 200, Minus-Eleven, Weymouth, MA). Environmental conditions are presented in **Table 2**. A total of 25 participants gave written informed consent to participate. Individuals with chronic health problems, illness at the time of testing, a history of exertional heat stroke or musculoskeletal injury were excluded from the studies. The University of Connecticut Institutional Review Board approved all study procedures.

Table 2 Environmental Conditions.

	Study A	Study B	Pooled
Ambient Temperature (°C)	40.1±1.0	39.5±1.9	39.8±1.7
Relative Humidity (%)	30.8±6.3	38.0±6.8	33.4±10.7
WBGT (°C)	30.8±1.6	31.9±2.3	31.1±1.9

Study A

11 male participants were enrolled in study A. Procedures consisted of eight trials that were performed in a randomized, counterbalanced, crossover design which evaluated athletic apparel beyond the scope of this investigation. Trials consisted of different t-shirt and hat ensembles that for the purposes of the present investigation can be considered similar.

Prior to these trials, subjects performed a familiarization trial to become acquainted with study procedures. At this time, we recorded a baseline body mass (BM) and height. In addition, we assessed estimated adiposity (body fat %, BF) using a 3-site skinfold method with Lange calipers (Beta Technology, Santa Cruz, CA).

Prior to each experimental trial, we provided participants instructions to consume 500mL of water at night before sleep and first thing in the morning. Before entering the environmental chamber participants inserted a rectal thermistor (Model 401, Measurement Specialties, Hampton, VA) 10cm past the anal sphincter to allow for the measurement of rectal temperature (T_{REC}). They also donned a heart rate (HR) monitor (ANT+ Heart Rate Monitor, Timex Group USA, Middlebury, CT) and researchers applied thermistors (Thermochron iButton, Embedded Data Systems, LLC., Lawrenceburg, KY) to seven sites on the skin surface: 1) chest (T_{Chest}), 2) abdomen (T_{abd}), 3) neck (T_{neck}), 4) back (T_{back}), 5) upper arm (T_{arm}), 6) thigh (T_{thigh}) and 7) calf (T_{calf}).

Participants entered the environmental chamber and sat quietly for 10 minutes to equilibrate. They then jogged on a motorized treadmill at a 5% grade between 7.2 and 9.6 km· hr^{-1} until they met one of the following stopping criteria: 1) volitional fatigue, 2) altered or uneven gait, 3) T_{RE} greater than 39.99°C, 4) HR greater than age predicted maximum (220-age) for 5 minutes or 5) 60 minutes of exercise. We recorded T_{REC}, HR, skin temperatures and environmental measures (WBGT, dry bulb temperature (T_{amb}), Relative Humidity (RH); Kestrel 4400, Nielsen-Kellerman Co., Boothwyn, PA) every 15 minutes.

Study B

14 participants (males, n=8; females, n=6) were enrolled in study B. Study procedures consisted of 3 identical exercise trials with 1 familiarization trial. During this familiarization trial, we recorded a baseline BM and height. In addition, we estimated adiposity using a 3-site skinfold method with Lange calipers (Beta Technology, Santa Cruz, CA).

Prior to each experimental trial, researchers provided participants instructions to consume 500mL of water at night before sleep and first thing in the morning. We tested female subjects during their luteal phase, based on a menstrual status history, to minimize the influence of the menstrual cycle on body temperature. Fluid consumption was restricted during all trials due to the presence of an esophageal thermometer.

Before entering the environmental chamber participants inserted a rectal thermometer (Model 401, Measurement Specialties, Hampton, VA) 10cm past the anal sphincter to allow for measurement of rectal temperature. They also donned a heart rate (HR) monitor (ANT+ Heart Rate Monitor, Timex Group USA, Middlebury, CT) and researchers applied thermistors (Thermochron iButton, Embedded Data Systems, LLC., Lawrenceburg, KY) to seven sites on the skin surface: 1) T_{Chest}, 2) T_{abd}, 3) T_{neck}, 4) T_{back}, 5) T_{arm}, 6) T_{thigh} and 7) T_{calf}. Participants entered the environmental chamber and sat quietly for 10 minutes to equilibrate with the hot environment (Table 2).

Participants then preformed a 20-minute exercise interval on a motorized treadmill. They alternated intervals of 5 minutes of walking at a 5% incline between 5.6 and 7.2 $km\cdot hr^{-1}$ and running at a 1% grade between 8.9 and 12.1 $km\cdot hr^{-1}$ until they met one of the following stopping criteria: 1) volitional fatigue, 2) altered or uneven gait, 3) T_{RE} greater than 39.99°C, 4) HR greater than age predicted maximum (220-age) for 5 minutes or 5) 60 minutes of exercise. We recorded T_{RE}, HR and skin temperatures every five minutes. We recorded environmental measures (WBGT, T_{amb}, RH; Kestrel 4400, Nielsen-Kellerman Co., Boothwyn, PA), every 15 minutes. Due to the frequency of measurement of environmental variables, data was assumed to not change significantly between measurement time points and were extended to the 5 minute intervals between actual recordings.

Statistical Analysis

Basic data analysis was performed in IBM SPSS Statistics 22 (IBM Corporation, Armonk, NY). Data are presented as mean±SD. First, we utilized Pearson's correlations to examine the relationship between predictors and T_{REC}. We then performed standard step-wise regression using all data cases from both studies for the basis of comparison with other techniques. The significance level was set *a priori* at $p<0.05$.

Model Development

For machine learning analysis, we utilized R version 3.2.3 (R Foundation for Statistical Computing, Vienna, Austria). Data cases from both studies were pooled and then randomly assigned to one of two groups: 1) Training (TRA) or 2) Testing (TEST). Inclusive of only complete data cases, TRA was allocated 75% of the data cases with the remaining 25% assigned to TEST.[57] This was performed in order to reduce over fitting models to the data set. Prediction models were developed based on TRA, while diagnostics were performed based on TEST. The data partitions were identical across all methods.

Two distinct approaches were used for machine learning analysis: regression and classification analysis. For all methods the following variables were entered into the models: 1) T_{chest}, 2) T_{abd}, 3) T_{neck}, 4) T_{back}, 5) T_{arm}, 6) T_{thigh}, 7) $T_{calf,}$, , 8) WBGT, 9) T_{amb}, 10) RH, 11) HR, 12) Sex, 13) Age, 14) Height, 15) BM, 16) BF, 17) Speed, and 18) Incline. Regression analysis used four methods to predict T_{RE}, given a set of predictors. We first performed linear regression onall predictors included in the analysis. Next we used recursive partitioning analysis, wherein heuristic analysis is used to create cut points for predictor variables resulting in nodes that meet certain qualifications for the predictors.[58] The predictive values for this analysis are the mean values for the model data set in each node. In a similar fashion we created an M5' pruned model tree that replaces the mean values at each terminal node with an individual multivariate regression equation.[59] Finally we utilized Multivariate Adaptive Regression splines which applies unique linear regression models over distinct points of the data.[60]

Classification analysis sought to identify individual data cases wherein T_{RE} was greater than 39°C. Data cases were labeled as "hyperthermic" or "not hyperthermic" and we ran them through two analyses. We first utilized K-nearest neighbor analysis which classified TEST cases based on the mean of 5 TRA cases with similar characteristics.[61] Finally, a C5.0 decision tree was constructed using recursive partitioning to classify cases.[62]

IV. Results

Participant demographics are presented in **Table 2**. We identified 527 complete data cases that were used for analysis. Pearson correlations between T_{REC} and predictor variables are shown in **Table 3**. T_{neck} (r=0.071, p=0.01), T_{thigh} (r=0.284, p<0.001), T_{calf} (r=0.484, p<0.001), HR (r=0.666, p<0.001), Incline (r=-0.278, p<0.001), WBGT (r=0.443, p<0.001), T_{amb} (r=0.283, p<0.001) and RH (r=0.257, p<0.001) were significantly correlated with T_{REC}.

Table 3 Participant Demographics

	Study A (n=11)	Study B (n=14)	Pooled (n=25)
Age (y)	24±5	24±3	24±4
Height (cm)	180±7	174±11	177±10
Body Mass (kg)	74.55±8.23	75.25±15.30	75.94±12.45
Body Fat (%)	9.83±2.35	19.23±5.70	15.31±6.55

Stepwise linear regression analysis revealed seven variables to exhibit a significant relationship with T_{REC} (R^2=0.612, SEE=0.52, p<0.001). A plot of the predicted values compared to the actual T_{REC} values is shown in **Figure 2**. The regression equation is shown below.

$$
\begin{aligned}
TREC = 28.096 + -0.116Tchest + 0.012Tabdom + 0.177Tneck + 0.086Tback \\
+ -0.052Tarm + 0.060Tthigh + 0.027Tcalf + 0.020HR - 0.096Speed \\
+ 0.034Incline + 0.072Tamb + 0.022RH - 0.024BF + 0.008BM \\
- 0.003Age - 0.020Height - 0.1422Sex
\end{aligned}
$$

Table 4 Pearson Correlations between predictors and T_{REC}.

Variable	T_{neck}	T_{thigh}	T_{calf}	HR	Incline	T_{amb}	RH
TREC	0.071*	0.284*	0.484*	0.666*	-0.278*	0.283*	0.257*

* Indicates significance at a 0.05 level.

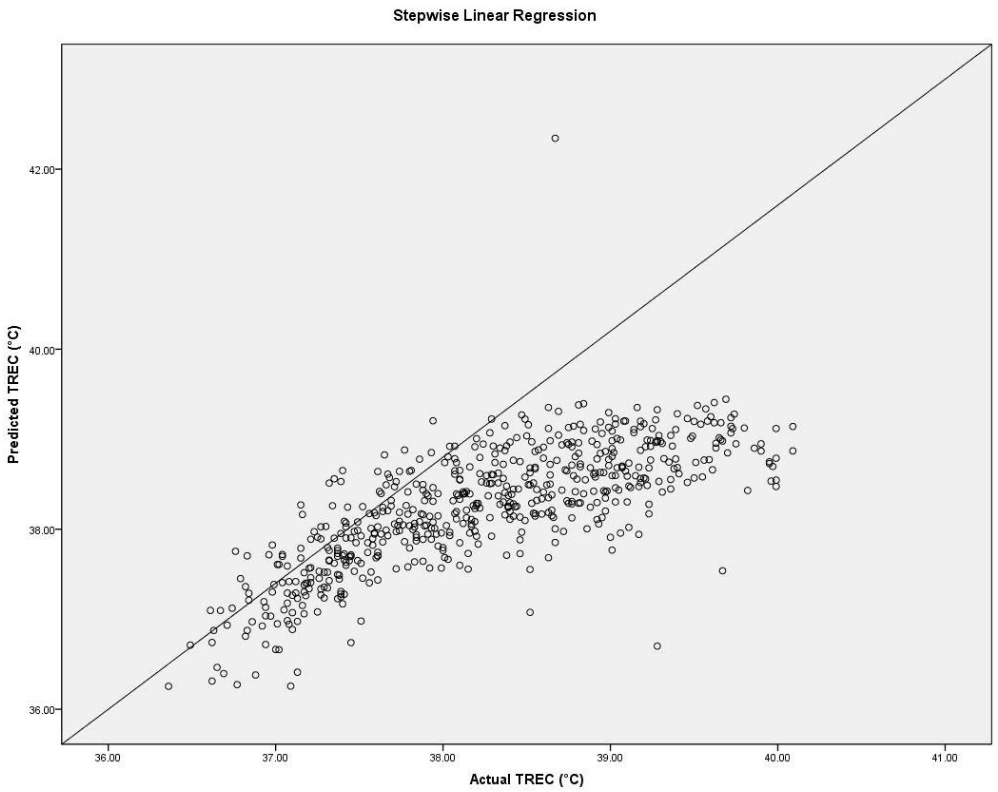

Figure 2. Predicted values from stepwise linear regression versus actual TEST values.

Machine Learning

Data partitioning yielded 396 cases in TRA and 131 cases in TEST. A comparison of regression models can be found in **Table 5**. Multivariate linear regression improved on stepwise regression (R^2=0.651, Root Mean Squared Error, RMSE=0.541°C); a plot of predicted values in comparison to actual TEST values can be found in **Figure 3**. The regression equation can be found below:

$$TREC = 27.075 + 0.020HR + 0.134Tneck + 0.066WBGT + 0.074Incline + 0.082Tthigh + -0.063Tarm + 0.004Humidity$$

Table 5. Comparison of Models.

Model	Predictors	R^2	Predicted TEST Values Correlation	Predicted TEST Values RMSE (°C)
Stepwise Linear Regression	HR, T_{neck}, WBGT, Incline, T_{thigh}, T_{arm}, RH	0.607	-	-
Multivariate Linear Regression	T_{Chest}, T_{abd}, T_{neck}, T_{back}, T_{arm}, T_{thigh}, $T_{calf,}$, WBGT, T_{amb}, RH, HR, Sex, Age, Height, BM, BF, Speed, and Incline	0.651	0.749	0.54
Recursive Partitioning	HR, T_{neck}, Height, T_{amb}, T_{calf}, Age	-	0.738	0.55
M5' Pruned Model Tree	HR, T_{back}, T_{neck}, RH, Height, T_{calf}, Age, T_{chest}, T_{thigh} Speed, T_{amb}, BF, BM	-	0.792	0.493
Multivariate Adaptive Regression Splines	T_{neck}, T_{arm}, T_{thigh}, T_{calf}, HR, Incline, RH, BF, BM, Height	0.776	0.848	0.428

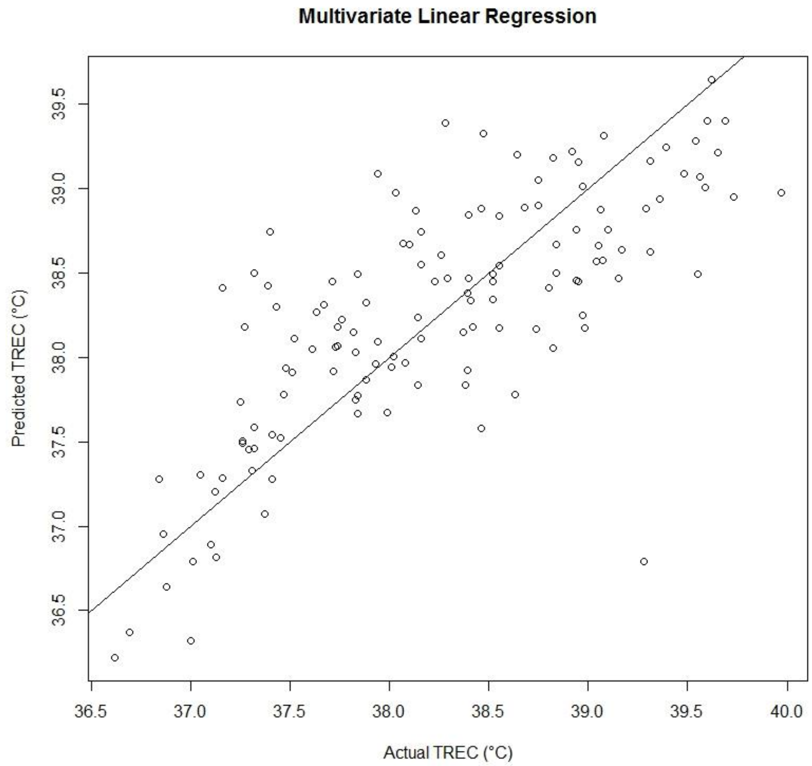

Figure 3 Predicted values from multivariate linear regression versus actual TEST values.

The results of a recursive partitioning analysis can be found in **Figure 4.** The relative performance of predicted values is plotted in **Figure 5**. Recursive partitioning yielded a similar RMSE as the previous regression models (RMSE=0.551°C).

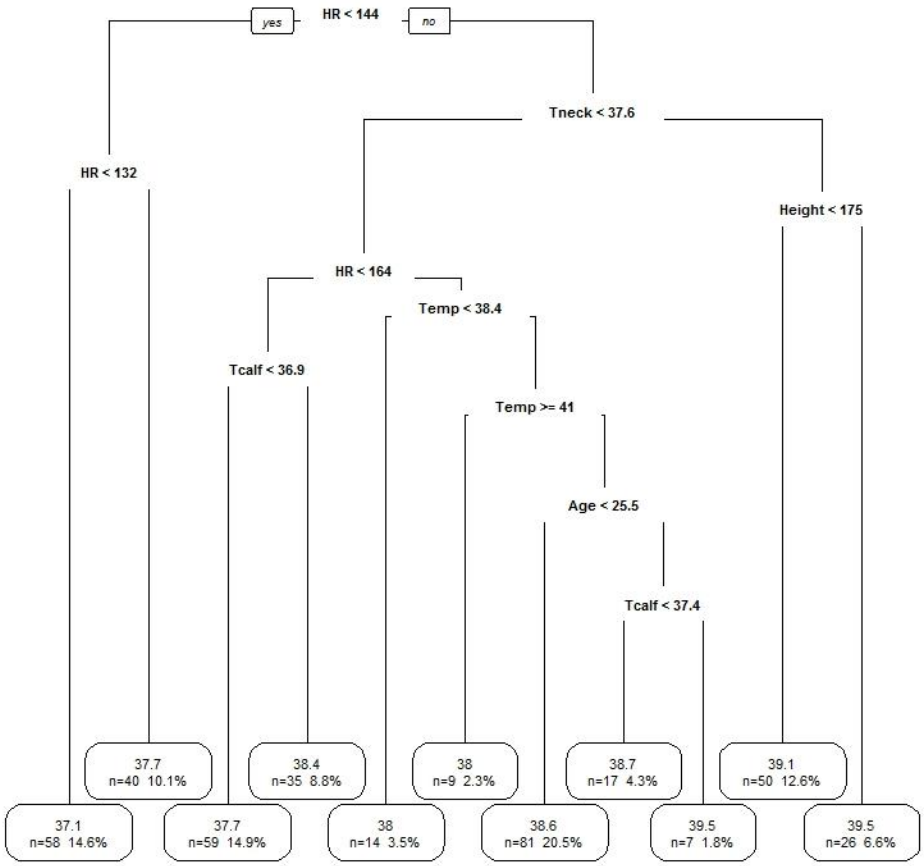

Figure 4 Recursive Partitioning Regression Tree. Each value represents a cut-point with two nodes: the left node is less than the cut-point and the right node is greater than the cut-point. In this way the tree dichotomizes data beginning with HR and ending with each terminal node based on the characteristics of that case.

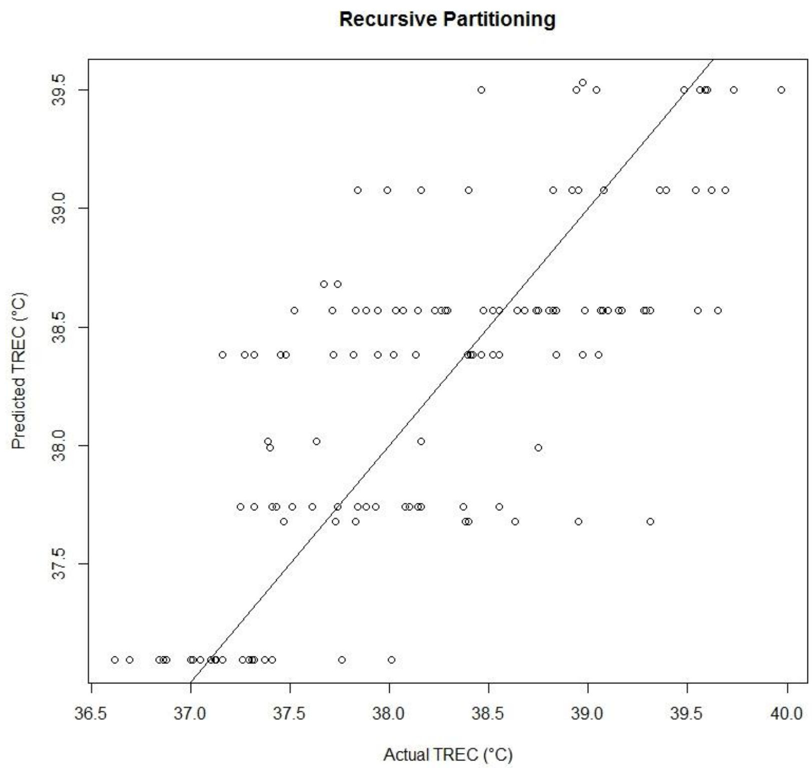

Figure 5. Predicted values from recursive partitioning versus actual TEST values.

A schematic of the M5' model tree is shown in **Figure 6**. The corresponding regression equations for each node can be found in **Table 6**. This model improved upon both the multivariate linear regression and recursive partitioning models (R^2=0.789, RMSE=0.493°C). A plot of the predicted values in contrast to actual TEST values is demonstrated in **Figure 7**.

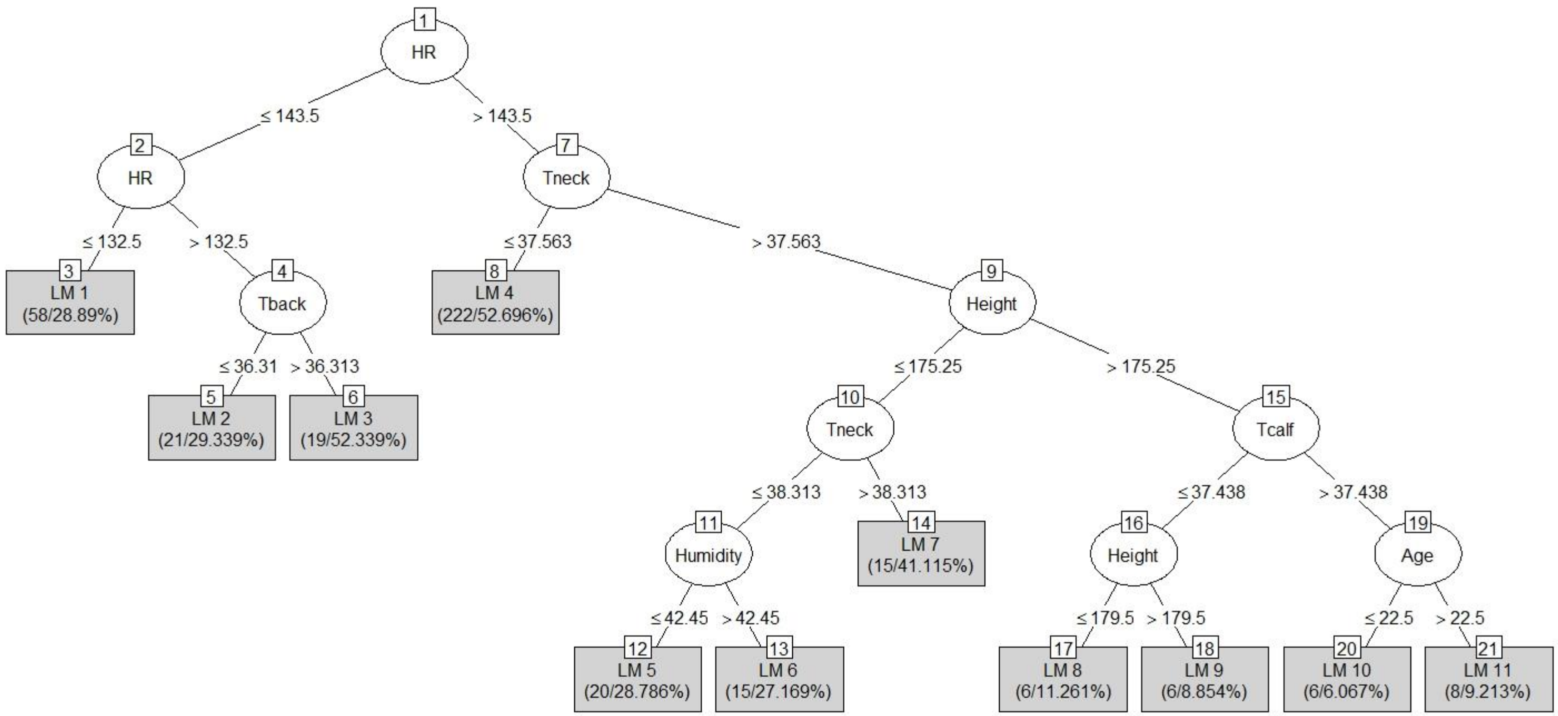

Figure 6. M5' Model Tree. Each value represents a cut-point with two nodes: the left node is less than the cut-point and the right node is greater than the cut-point. In this way the tree dichotomizes data beginning with HR and ending with each terminal node based on the characteristics of that case.

Table 5 M5' Model Equations.

Node	Equation
LM1	TREC = -0.018Tchest + 0.022Tneck + 0.030Tback + 0.092Tthigh + 0.011HR - 0.218Speed + 0.009Tamb + 0.006RH - 0.004BF + 0.001BM - 0.002Height + 32.169
LM2	TREC = 0.05Tchest - 0.043Tabdom + 0.062Tneck + 0.036Tback + 0.036Tthigh + 0.064Tcalf + 0.006HR - 0.019Speed + 0.009Tamb + 0.008RH - 0.004 B F + 0.001 * BM - 0.002Height + 29.147
LM3	TREC = 0.054Tchest - 0.045Tabdom + 0.064Tneck + 0.037Tback + 0.037 Tthigh - 0.03Tcalf + 0.006HR - 0.168Speed + 0.009Tamb + 0.007RH - 0.004B F + 0.001BM - 0.002Height + 33.171
LM4	TREC = -0.127Tchest + 0.071Tneck + 0.109Tback - 0.075Tarm + 0.068Tthigh + 0.191Tcalf + 0.018HR - 0.01Speed + 0.049Tamb + 0.018RH - 0.039BF + 0.007BM+ 0.012Age - 0.024Height + 28.526
LM5	TREC = -0.075Tchest + 0.026Tabdom + 0.116Tneck + 0.022Tback - 0.05Tthigh + 0.063Tcalf + 0.008HR - 0.029Speed + 0.119Tamb + 0.022RH - 0.008 BF + 0.001Height + 28.226
LM6	TREC = -0.075Tchest + 0.026Tabdom+ 0.259Tneck + 0.022Tback- 0.054Tthigh + 0.067Tcalf + 0.007HR - 0.029Speed + 0.124Tamb + 0.005RH - 0.008 BF + 0.001Height + 23.506
LM7	TREC = -0.075Tchest + 0.026Tabdom + 0.138Tneck + 0.022Tback - 0.045Tthigh + 0.054Tcalf + 0.007HR - 0.029Speed + 0.112Tamb + 0.005RH - 0.008 BF + 0.001Height + 28.966
LM8	TREC = -0.175Tchest + 0.041Tabdom + 0.224Tneck + 0.139Tback - 0.001Tthigh + 0.063Tcalf + 0.005HR - 0.029Speed + 0.121Tamb + 0.005RH - 0.007 5BF - 0.020Height + 26.22
LM9	TREC = -0.175Tchest + 0.041Tabdom + 0.242Tneck + 0.116Tback - 0.001Tthigh + 0.063Tcalf + 0.005HR - 0.029Speed + 0.121Tamb + 0.005RH - 0.008 BF - 0.02Height + 26.329
LM10	TREC = -0.170Tchest + 0.041Tabdom + 0.179Tneck + 0.085Tback - 0.001Tthigh + 0.060Tcalf + 0.005HR - 0.029Speed+ 0.118Tamb + 0.005RH- 0.008 BF - 0.025Age- 0.01Height + 28.928
LM11	TREC = -0.169Tchest + 0.041Tabdom + 0.179Tneck + 0.085Tback - 0.001Tthigh + 0.060Tcalf + 0.005HR - 0.029Speed + 0.081Tamb + 0.005RH - 0.007 BF - 0.023Age - 0.010Height + 30.321

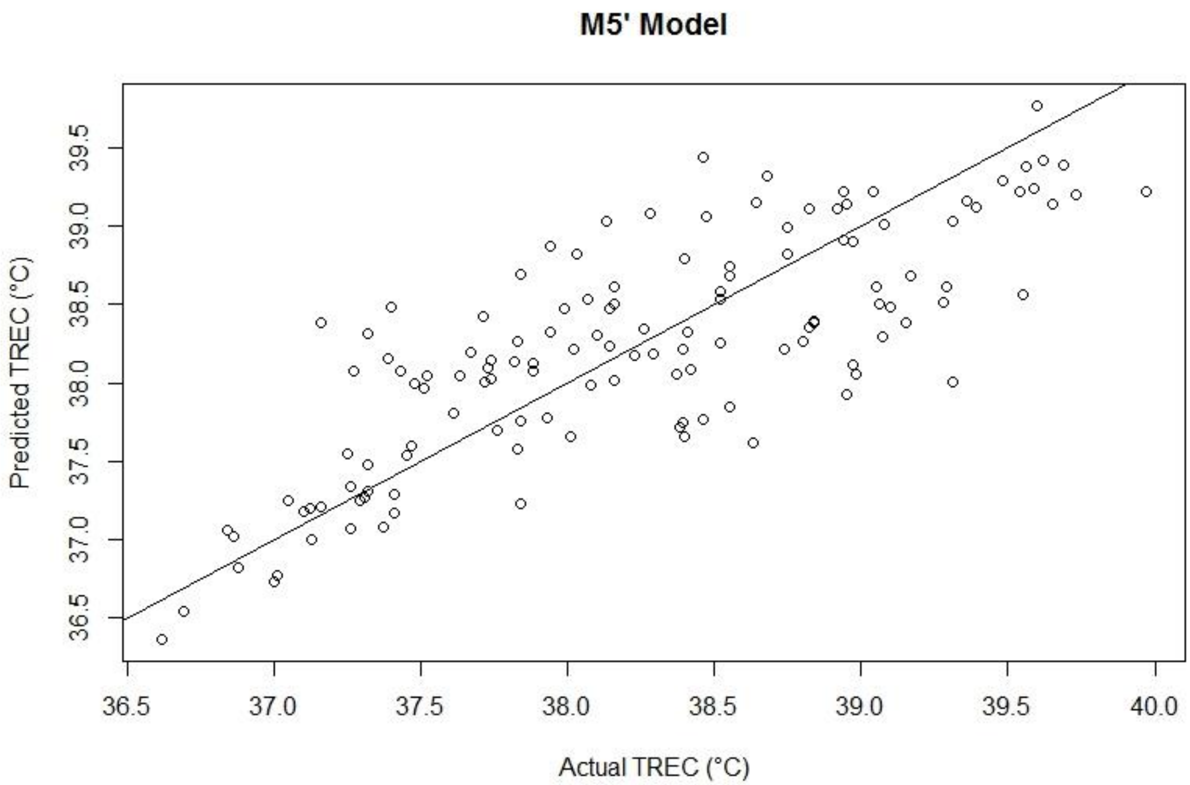

Figure 7. Predicted values from M5' model tree versus actual TEST values.

The final regression model, multivariate adaptive regression splines, demonstrated the greatest predictive capabilities. **Figure 8** plots the performance of this model on TEST data sets. Multivariate adaptive regression splines demonstrated the greatest predictive ability of the regression models with a RMSE of 0.428°C, despite a smaller R^2 (0.776).

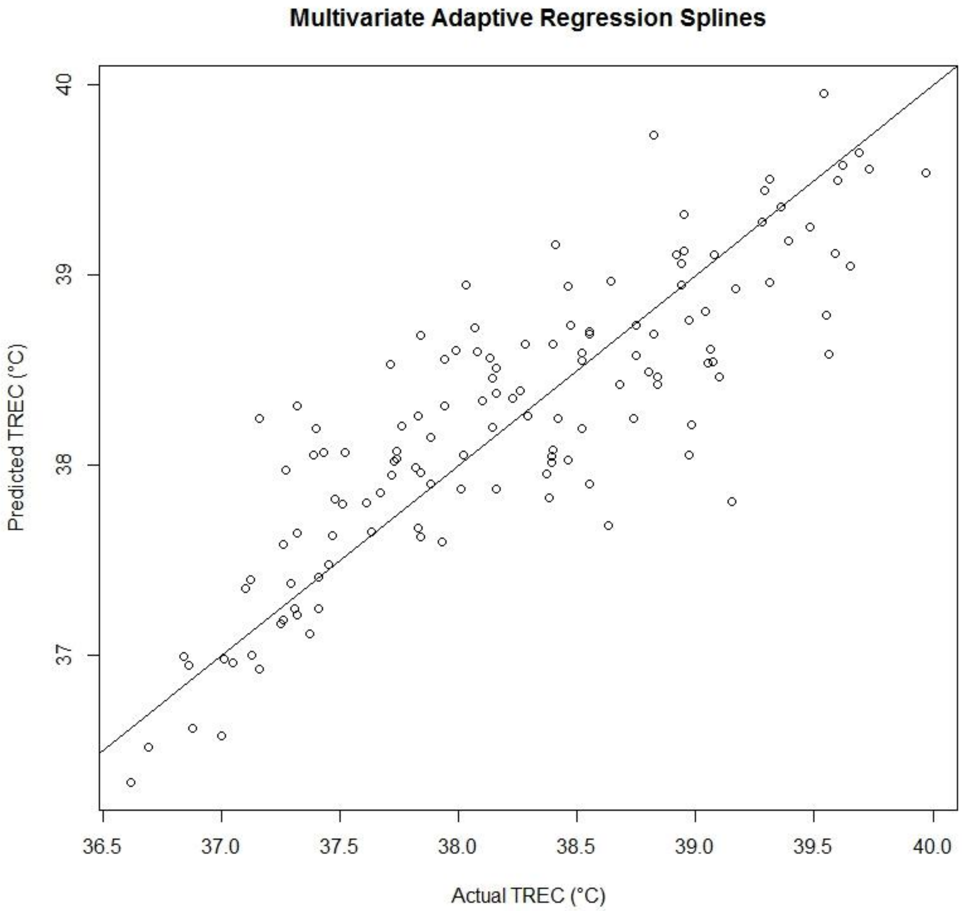

Figure 8. Predicted values from multivariate adaptive regression splines versus actual TEST values.

The results of a K-nearest neighbor analysis are shown in **Table 7**. Using TEST data, a sensitivity of 0.55 and a specificity of 0.87 were calculated. From there we calculated a positive likelihood ratio of 4.23 and a negative likelihood ratio of 0.51.

Table 6. K-nearest neighbor confusion matrix.

		Predicted		
		TREC >39°C	TREC <39°C	Total
Actual	TREC >39°C	11	14	25
	TREC <39°C	9	97	106
	Total	20	111	131

The C5.0 decision tree is shown in **Figure 9**. A confusion matrix of these results applies to TEST data are presented in **Table 8**. We calculated a sensitivity of 0.625 and a specificity of 0.906 for this model. This in turn yielded a positive likelihood ratio of 6.58 and a negative likelihood ratio of 0.413.

Table 7 C5.0 Confusion Matrix.

		Predicted		
		TREC >39°C	TREC <39°C	Total
Actual	TREC >39°C	15	10	25
	TREC <39°C	9	97	106
	Total	24	107	131

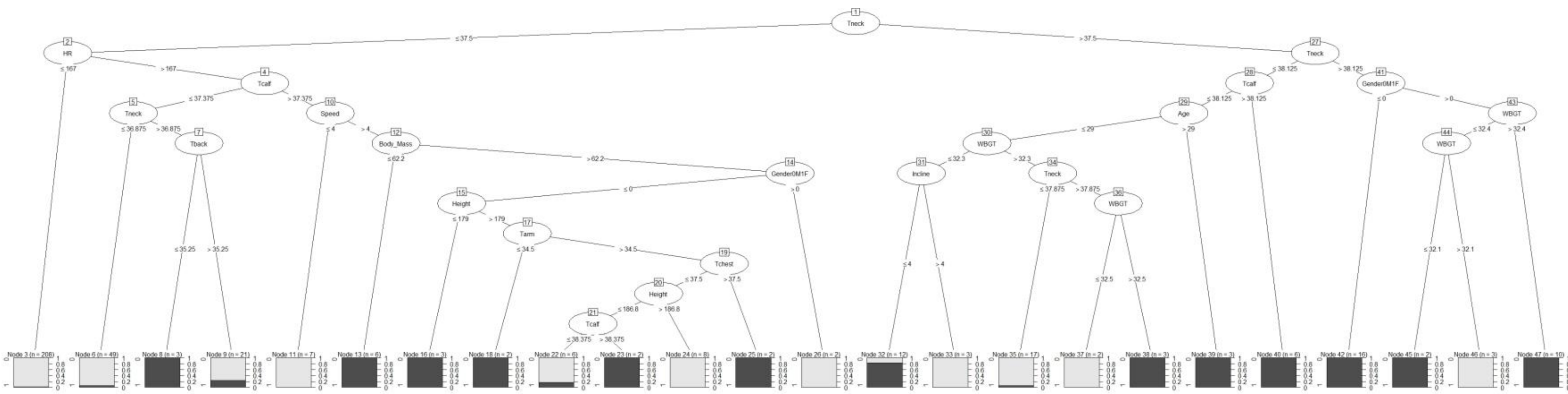

Figure 9 C5.0 Decision Tree. Each value represents a cut-point with two nodes: the left node is less than the cut-point and the right node is greater than the cut-point. In this way the tree dichotomizes data beginning with Tneck and ending with each terminal node based on the characteristics of that case.

V. Discussion

In this study we presented the results of using non-invasive physiological measures to estimate internal body temperature during strenuous exercise in the heat. A comparison of our Multivariate Adaptive Regression Splines model to previous research is presented in Table 9. Previous studies in similar populations report RMSE values of 0.2 to 0.34°C. [46,47,55] Our best model, a multivariate adaptive regression splines equation, resulted in a RMSE of 0.428°C. In a similar fashion our classification models were able to generate predictions of whether an individual was hyperthermic with a specificity of 0.906 and a sensitivity of 0.625. Depending on the use case these models can provide some relevant information.

Through several machine learning techniques, we demonstrate improvements in the estimation of body temperature from non-invasive measures compared to a standard linear regression analysis. As shown in Figure 2, a standard stepwise linear regression demonstrates non-uniformity across the range of data. This indicates that a non-linear or alternative approach is necessary. Our machine learning models, in particular the multivariate adaptive regression model, were more capable of capturing this non-uniformity in the data.

Table 9. Comparison of models to predict body temperature during exercise in the heat.

Study	Environmental Conditions	Exercise Protocol	Variables	Modeling Technique	Diagnostic Outcome
Xu et al., 2013	25°C, 50%RH; 35°C, 70%RH; 42°C, 25%RH; Army combat uniform with body armor; Laboratory	2h treadmill walking at 350W and 540W	Sternum T_{sk}, Sternum Heat Flux	Linear Regression	R^2=0.75
Niedermann et al., 2013	10°C, 30°C; Laboratory	Treadmill running 40% and 60% VO_{2peak}	HR, Chest Heat Flux, Back Heat Flux, Upper Arm T_{sk}, Lower Arm T_{sk}, Thigh T_{sk}	Principle Component and Linear Regression	RMSE=0.28-0.34°C
Buller et al., 2013	24-35°C, 42-97%RH; Army combat uniform with body armor; Outdoors	24h military field exercise	Heart Rate	Kalman Filter	Bias= -0.003±0.32, RMSE= 0.30±0.13
Kim et al., 2015	29.5 to 25.5°C; firefighter PPE; Laboratory	60 minutes of treadmill walking	Chest and Forehead T_{sk}	Linear Regression	T_{chest}, R^2= 0.826; $T_{forehead}$, R^2= 0.824
Richmond et al., 2015	25°C, 50%RH; 35°C, 35%RH; 40°C, 25%RH; variety of clothing conditions; Laboratory	40 minutes of walking with 20 minutes of rest	Insulated 11-site T_{sk}, microclimate T_{sk}, HR, and work	Bootstrap Regression	R^2= 0.86, SEE=0.27°C
Belval et al., 2016	39.8±1.7°C, 33.4±10.7%RH; Laboratory	Treadmill walking and jogging	T_{neck}, T_{arm}, T_{thigh}, T_{calf}, HR, Incline, RH, BF, BM, Height	Multivariate Adaptive Linear Regression Splines	R^2=0.776 RMSE= 0.428°C

In our models we were able to increase the R^2 values from 0.607 for a stepwise regression to 0.776 for a multivariate adaptive regression splines model. Likewise, the more advanced model resulted in a decrease of 0.093°C in RMSE compared to the traditional model, however the RMSE for stepwise regression is based on all data points. Interestingly, recursive partitioning, a more easily implemented model, performed similarly to stepwise and multivariate linear regression models (Multivariate Linear Regression, RMSE=0.54°C; Recursive Partitioning (0.55°C), although **Figure 5** demonstrates this stratification can widely over- or under-estimate individual values.

In contrast to previously published studies, our results illustrate some of the challenges to modeling body temperature during strenuous exercise in the heat. Buller et al. was able to a construct a model that predicted internal temperature with a RMSE of 0.30±0.13°C using only a series of heart rate measures.[55] However, the construct of their experiments differs from the present investigation. For example, Buller et al. utilized a Kalman filter, which is a recursive function that relies on a time series of data to predict values whereas our models relied only on one-time point to generate a prediction.

In addition, their model was developed based upon a 24-hour military training exercise with body armor and other protective equipment worn by soldiers, while our study utilized individuals running at different intensities in a laboratory environment. A possible explanation for the difference in diagnostic accuracy may be that predictor variables are inherently more variable during intense exercise. For example, in our experiment cardiac drift could have caused more variability in heart rate measures independent of body temperature compared to a prolonged military marching scenario. Casa et al. demonstrated that even small changes in hydration can have an impact on heart rate during exercise in the heat.[36]

Kim and Lee also used skin temperature to predict deep-body temperature.[48] Since they undertook their study in firefighters wearing personal protective equipment, they found much higher associations between skin temperature and internal body temperature. They found high Pearson product correlations for forehead (r=0.908) and chest (r=0.908) skin temperatures whereas our highest skin temperature correlation was for the calf (r=0.484). This likely is due to the microclimate created by firefighter's equipment which likely impeded heat loss.

Pandolf and Goldman demonstrated in 1978 that a convergence of skin temperature and internal body temperature represented the upper limits of human thermoregulatory ability.[63] The difference between our results and those by Kim and Lee indicate that the extent of this phenomenon may differ depending on the nature of the exertional hyperthermia. In our study, skin temperature, which is greatly influenced by the skin blood flow, did not on its own predict high rectal temperatures (R^2=0.287). Therefore, it appears that internal body temperature increases were not impeded by a diminished heat loss.

Niedermann et al. conducted the experiment most similar to the present investigation.[46] In their protocol participants alternated running between 40% and 60% of VO_{2peak} in both 10°C and 30°C environments. Their principle component analysis relied on three skin temperature sites, two skin heat flux sensors and heart rate; this yielded a RMSE of 0.28°C to 0.34°C across validation data sets. The use of heat flux could be a potential additional variable to consider in future models.[64]

The investigation by Richmond et al. has the lowest reported RMSE for a prediction model at 0.27°C.[47] However, the practical implementation of this technique may be difficult as it relies on 11 skin temperature site and microclimate temperatures within clothing. This study also examined classification of hyperthermic individuals and found a remarkable sensitivity of 0.97 and a specificity of 0.86. This contrasts with our models which found higher specificities (0.906) but much lower sensitivities (0.625). Once again, as Richmond et al. used a walking protocol it is possible that intense exercise introduces additional variation.

While our regression models did not meet the limits set forth by Moran and Mendal of 0.1°C,[65] several other standards exist that may be more applicable. For example, Gunga et al. proposed an acceptable difference of 0.5°C that may be more applicable in situations where a model or device are not the ultimate diagnostic tool.[64] In the same way, Buller et al. reported their acceptable difference as 0.4°C, as this is the difference between measurement of esophageal and rectal temperature.[55] Our multivariate adaptive regression splines model approaches this value (0.428°C).

Arguably, the largest use case for prediction models of internal body temperature would be to use them for the screening of individuals who may need to either rest or be evaluated by medical staff during exercise in the heat. For this reason, we utilized classification methods to identify whether or not an individual would be considered hyperthermic, >39°C. With a positive likelihood ratio of 6.58, our C5.0 model demonstrates good positive predictive value. In other words, an individual who meets the criteria of our model is 6.58 times more likely to be hyperthermic than someone who does not. While this is mirrored by a less strong negative likelihood ratio, the use case of these models are supportive of our results. A clinician would likely rather evaluate more individuals who are potentially hyperthermic than miss some individuals who are not captured by the model.

Limitations

The largest limitation of this study was the nature of the data set utilized for model development. In comparison to machine learning datasets used in other disciplines our sample size was small. It is possible, that with a larger sample, prediction equations could be strengthened, especially since the present investigation demonstrated high R^2 values but moderate RMSE. Furthermore, both studies included in the present investigation utilized similar exercise protocols in similar environments. In order to increase the external validity of models, a wide variety of situations with a diverse study population should be tested. While we utilized some female participants, they were only tested in their luteal phase, with known variation in body temperature occurring throughout the menstrual cycle.[66]

In addition, our models only apply to the circumstances captured within the confines of our experiments. All models developed for the use of prediction of internal body temperature need to be validated in a variety of settings prior to use. For example, our model needs to be validated in real-world athletic situations where both exercise intensity and environmental conditions are very dynamic compared to a laboratory. Furthermore, no model has been developed to account for the effects of heat acclimatization, a process that is well known to impact thermoregulation.

Relative model performance can also be attributed to predictor selection. In our investigation we sought to identify predictors that could be easily measured or known by a lay population. However, variables excluded from analysis could add significant predictive value.

For example, urine hydration measures were not included in our models, both due to the difficulty in measurement and the imperfect agreement as measures of hydration. However, our current understanding of exercise in the heat is that hydration plays a fundamental role in the limits of our physiology and therefore would have value in a prediction model.

Finally, this was not an investigation of exertional heat illnesses. While increased body temperature is strongly linked with a host of illnesses, the temperature wherein an individual succumbs to a heat illness is highly individual.[4] Even patients with exertional heat stroke present with temperatures ranging from 40°C to 44°C.[12] Therefore, there are clearly other factors should be considered if future models are developed to predict the onset of exertional heat illnesses. Future research into prediction models should focus on two areas: 1) improving existing models utilizing new techniques and 2) identification of alternative predictors. The models presented within this investigation are well documented and validated within the field of machine learning.[57-62] Therefore, newer or alternative combinations of techniques may present greater opportunities. For example, the incorporation of a Kalman filter or other recursive sampling technique could greatly improve the predictive abilities of the models presented here, albeit increase the technological burden for application in a field setting.

As new technologies come to light, such as non-invasive hydration measurement, future prediction models should incorporate the best measurement of the underlying physiology they can obtain. For example, skin temperature measurements via thermal imaging could represent a better portrayal of total skin temperature than a thermocouple approach.[67] It is fundamentally important that future models be based on sound physiological reasoning for the inclusion of predictors, rather than trying to model noisy variables utilizing advanced techniques.

With the advent of models that are able to estimate deep body temperature within acceptable limits, it may be possible to reconsider the manner in which body temperature measurements are used within an athletic or occupational setting. We suggest that body temperature has the potential to be used as more than a medical diagnostic tool, but also as a metric for both occupational effectiveness and athletic training. Easily accessible estimates of body temperature could help supervisors and coaches assess the responses of individuals to environmental stressors more easily, allowing for more individualized training and acclimatization.

Conclusion

In conclusion, we present several different machine learning methods to utilize in the development of prediction models for internal body temperature during exercise in the heat. For a regression model, we found a multivariate adaptive regression splines model performed best. Meanwhile, a C5.0 decision tree model was found to have good positive predictive value of whether or not an individual is hyperthermic. Although these are not a substitution for direct measurement of body temperature in the case of a suspected exertional heat illness, they represent a tool that clinicians could potentially use to assist athletes, laborers and soldiers exercising in the heat. Future research is required to refine prediction models for internal body temperature during exercise in the heat.

VI. References

1. Heinrich B. Why we run: a natural history. 2002. HaprersCollins, New York, NY.

2. Nadel E, Wenger CB, Roberts MF, Stolwijk JA, Cafarelli E. Physiological Defenses against Hyperthermia of Exercise. *Annals of the New York Academy of Sciences*. 1977;301(1):98-109.

3. Nybo L. Brain temperature and exercise performance. *Experimental Physiology*. 2012;97(3):333-339.

4. Casa DJ, Demartini JK, Bergeron MF, et al. National Athletic Trainers' Association Position Statement: Exertional Heat Illnesses. *Journal of Athletic Training*. 2015;50(9):986-1000.

5. Armstrong LE. Exertional Heat Illnesses. 2003:275. Human Kinetics, Champaign, IL.

6. Armstrong LE, Anderson JM. Heat Exhaustion, Exercise-Associated Collapse, and Heat Syncope. *Exertional Heat Illnesses*. 2003. Human Kinetics, Champaign, IL.

7. Leon LR, Helwig BG. Heat stroke: Role of the systemic inflammatory response. *Journal of Applied Physiology*. 2010;109(6):1980-1988.

8. Casa DJ, Armstrong LE, Kenny GP, O'Connor FG, Huggins RA. Exertional heat stroke: new concepts regarding cause and care. *Current Sports Medicine Reports*. 2012;11(3):115-123.

9. Allen SB, Cross KP. Out of the frying pan, into the fire: a case of heat shock and its fatal complications. *Pediatric emergency care*. 2014;30(12):904-910.

10. Rav-Acha M, Hadad E, Epstein Y, Heled Y, Moran DS. Fatal exertional heat stroke: a case series. *The American Journal of the Medical Sciences*. 2004;328(2):84-87.

11. Boden BP, Breit I, Beachler JA, Williams A, Mueller FO. Fatalities in High School and College Football Players. *The American Journal of Sports Medicine*. 2013;41(5):1108-1116.

12. Demartini JK, Casa DJ, Stearns R, et al. Effectiveness of cold water immersion in the treatment of exertional heat stroke at the falmouth road race. *Medicine & Science in Sports & Exercise*. 2015;47(2):240-245.

13. Armed Forces Health Surveillance Center (AFHSC). Update: Heat injuries, active component, U.S. Armed Forces, 2013. *MSMR*. 2014;21(3):10-13.

14. Casa DJ, Armstrong LE, Ganio MS, Yeargin SW. Exertional heat stroke in competitive athletes. *Current Sports Medicine Reports*. 2005;4(6):309-317.

15. Casa DJ, McDermott BP, Lee EC, Yeargin SW, Armstrong LE, Maresh CM. Cold water immersion: the gold standard for exertional heatstroke treatment. *Exercise and Sport Sciences Reviews*. 2007;35(3):141-149.

16. Sithinamsuwan P, Piyavechviratana K, Kitthaweesin T, et al. Exertional Heatstroke: Early Recognition and Outcome With Aggressive Combined Cooling—a 12-Year Experience. *Military Medicine*. 2009;174(5):496-502.

17. Casa DJ. Preventing Sudden Death in Sport and Physical Activity. 2011:367.

18. Kerr ZY, Marshall SW, Comstock RD, Casa DJ. Exertional heat stroke management strategies in United States high school football. *The American Journal of Sports Medicine*. 2014;42(1):70-77.

19. Castle PC, Maxwell N, Allchorn A, Mauger AR, White DK. Deception of ambient and body core temperature improves self paced cycling in hot, humid conditions. *European Journal of Applied Physiology*. 2011;112(1):377-385.

20. Wingo JE, Ganio MS, Cureton KJ. Cardiovascular drift during heat stress: implications for exercise prescription. *Exercise and Sport Sciences Reviews*. 2012;40(2):88-94.

21. Wingo JE, Cureton KJ. Body cooling attenuates the decrease in maximal oxygen uptake associated with cardiovascular drift during heat stress. *European Journal of Applied Physiology*. 2006;98(1):97-104.

22. Casa DJ. Exercise in the heat. I. Fundamentals of thermal physiology, performance implications, and dehydration. *Journal of Athletic Training*. 1999;34(3):246.

23. Stitt J. Central Regulation of Body Temperature. 2001;6(1):1-39.

24. Givoni B, Goldman RF. Predicting rectal temperature response to work, environment, and clothing. *Journal of Applied Physiology*. 1972;32(6):812-822.

25. Selkirk GA, McLellan TM. Influence of aerobic fitness and body fatness on tolerance to uncompensable heat stress. *Journal of Applied Physiology*. 2001;91(5):2055-2063.

26. Cheung S. Advanced environmental exercise physiology. 2010. Human Kinetics, Champaign, IL.

27. Cramer MN, Morris NB, Jay O. Dissociating biophysical and training-related determinants of core temperature. *Exercise and Sport Sciences Reviews*. 2012;40(3):183.

28. Deren TM, Coris EE, Casa DJ, et al. Maximum heat loss potential is lower in football linemen during an NCAA summer training camp because of lower self-generated air flow. *The Journal of Strength & Conditioning Research*. 2014;28(6):1656-1663.

29. Dervis S, Coombs GB, Chaseling GK, Filingeri D, Smoljanić J, Jay O. A comparison of thermoregulatory responses to exercise between mass-matched groups with large differences in body fat. *Journal of Applied Physiology*. 2016;120(6):615-623.

30. Cramer MN, Jay O. Explained variance in the thermoregulatory responses to exercise: the independent roles of biophysical and fitness/fatness-related factors. *Journal of Applied Physiology*. 2015;119(9):982-989.

31. Grundstein AJ, Ramseyer C, Zhao F, et al. A retrospective analysis of American football hyperthermia deaths in the United States. *International journal of biometeorology*. 2012;56(1):11-20.

32. Budd GM. Wet-bulb globe temperature (WBGT)—its history and its limitations. *Journal of Science and Medicine in Sport*. 2008;11(1):20-32.

33. Rowlinson S, Rowlinson S, Yunyanjia A, et al. Management of climatic heat stress risk in construction: a review of practices, methodologies, and future research. *Accident; analysis and prevention*. 2014;66:187-198.

34. Gagnon D, Jay O, Kenny GP. The evaporative requirement for heat balance determines whole-body sweat rate during exercise under conditions permitting full evaporation. *The Journal of Physiology*. 2013;591:2925-2935.

35. Demartini JK, Casa DJ, Belval LN, et al. Environmental conditions and the occurrence of exertional heat illnesses and exertional heat stroke at the Falmouth Road Race. *Journal of Athletic Training*. 2014;49(4):478-485.

36. Casa DJ, Stearns RL, Lopez RM, et al. Influence of hydration on physiological function and performance during trail running in the heat. *Journal of Athletic Training*. 2010;45(2):147-156.

37. Armstrong LE, Johnson EC, Casa DJ, et al. The American football uniform: uncompensable heat stress and hyperthermic exhaustion. *Journal of Athletic Training*. 2010;45(2):117-127.

38. Casa DJ, Becker SM, Ganio MS, et al. Validity of devices that assess body temperature during outdoor exercise in the heat. *Journal of Athletic Training*. 2007;42(3):333.

39. Saltin B, Hermansen L. Esophageal, rectal, and muscle temperature during exercise. *Journal of Applied Physiology*. 1966;21(6):1757-1762.

40. Mazerolle SM, Scruggs IC, Casa DJ, et al. Current Knowledge, Attitudes, and Practices of Certified Athletic Trainers Regarding Recognition and Treatment of Exertional Heat Stroke. *Journal of Athletic Training*. 2010;45(2):170-180.

41. Mazerolle SM, Ganio MS, Casa DJ, Vingren J, Klau JF. Is oral temperature an accurate measurement of deep body temperature? A systematic review. *Journal of Athletic Training*. 2011;46(5):566-573.

42. Ganio MS, Brown CM, Casa DJ, et al. Validity and reliability of devices that assess body temperature during indoor exercise in the heat. *Journal of Athletic Training*. 2009;44(2):124-135.

43. Huggins R, Glaviano N, Negishi N, Casa DJ, Hertel J. Comparison of Rectal and Aural Core Body Temperature Thermometry in Hyperthermic, Exercising Individuals: A Meta-Analysis. *Journal of Athletic Training*. 2012;47(3):329-338.

44. Byrne C, Lim CL. The ingestible telemetric body core temperature sensor: a review of validity and exercise applications. *British Journal of Sports Medicine*. 2007;41(3):126-133.

45. Moran D, Shapiro Y, Epstein Y, Burstein R. Validation and adjustment of the mathematical prediction model for human rectal temperature responses to outdoor environmental conditions. *Ergonomics*. 1995.

46. Niedermann R, Wyss E, Annaheim S, Psikuta A, Davey S, Rossi RM. Prediction of human core body temperature using non-invasive measurement methods. *International journal of biometeorology*. 2013;58(1):7-15.

47. Richmond VL, Davey S, Griggs K, Havenith G. Prediction of Core Body Temperature from Multiple Variables. *The Annals of occupational hygiene*. 2015;59(9):1168-1178.

48. Kim S, Lee J-Y. Skin sites to predict deep-body temperature while wearing firefighters' personal protective equipment during periodical changes in air temperature. *Ergonomics*. 2015:1-8.

49. Buller MJ, Tharion WJ, Cheuvront SN, et al. Estimation of human core temperature from sequential heart rate observations. *Physiological measurement*. 2013;34(7):781-798.

50. Buller MJ, Welles AP, Stower J, et al. Thermal-Work Strain during Marine Rifle Squad Operations in Afghanistan (March 2010). *Military Medicine*. 2011;178(10):1141-1148.

51. Schlader ZJ, Perry BG, Jusoh MRC, Hodges LD, Stannard SR, Mundel T. Human temperature regulation when given the opportunity to behave. *European Journal of Applied Physiology*. 2013;113(5):1291-1301.

52. Hayes M, Castle PC, Ross EZ, Maxwell NS. The Influence of Hot Humid and Hot Dry Environments on Intermittent-Sprint Exercise Performance. *International journal of sports physiology and performance*. 2014;9(3):387-396.

53. Pryor RR, Seitz JR, Morley J, et al. Estimating Core Temperature with External Devices After Exertional Heat Stress in Thermal Protective Clothing. *Prehospital Emergency Care*. 2012;16(1):136-141.

54. Buller MJ, Tharion WJ, Cheuvront SN, et al. Estimation of human core temperature from sequential heart rate observations. *Physiological measurement*. 2013;34(7):781-798.

55. Buller MJ, Tharion WJ, Cheuvront SN, et al. Estimation of human core temperature from sequential heart rate observations. *Physiological measurement*. 2013;34(7):781-798.

56. Richmond VL, Davey S, Griggs K, Havenith G. Prediction of Core Body Temperature from Multiple Variables. *The Annals of occupational hygiene*. 2015;59(9):1168-1178.

57. Kuhn M, Wing J, Weston S, Williams A, Keefer C. caret: classification and regression training. R package version 6.0-24. 2014.

58. Therneau T, Atkinson B, Ripley B. Recursive partitioning and regression trees. 2015.

59. Hornik K, Buchta C, Zeileis A. Open-source machine learning: R meets Weka. *Computational Statistics*. 2008;24(2):225-232.

60. Friedman JH. Multivariate adaptive regression splines. *The annals of statistics*. 1991.

61. Ripley BD. Pattern Recognition and Neural Networks. *Cambridge*. 1996.

62. Quinlan JR. C4.5: Programs for empirical learning. 1993.

63. Pandolf KB, Goldman RF. Convergence of skin and rectal temperatures as a criterion for heat tolerance. *Aviation, Space, and Environmental Medicine*. 1978;49:1095-1101.

64. Gunga HC, Werner A, Stahn A, et al. The Double Sensor-A non-invasive device to continuously monitor core temperature in humans on earth and in space. *Respiratory Physiology & Neurobiology*. 2009;169 Suppl 1:NaN-NaN.

65. Moran DS, Mendal L. Core temperature measurement: methods and current insights. *Sports Medicine*. 2002;32(14):879-885.

66. Kazman JB, Purvis DL, Heled Y, et al. Women and Exertional Heat Illness: Identification of Gender Specific Risk Factors. *US Army Medical Department journal*. 2015:48-56.

67. Choi JK, Miki K, Sagawa S, Shiraki K. Evaluation of mean skin temperature formulas by infrared thermography. *International journal of biometeorology*. 1997;41(2):68-75.