Annals of Work Exposures and Health, 2017, Vol. 61, No. 6, 621–632
doi: 10.1093/annweh/wxx035
Advance Access publication 8 June 2017
Original Article

Original Article

Ability to Discriminate Between Sustainable and Unsustainable Heat Stress Exposures—Part 2: Physiological Indicators

Ximena P. Garzón-Villalba, Yougui Wu, Candi D. Ashley, Thomas E. Bernard*

College of Public Health, University of South Florida, 13201 Bruce B Downs Blvd, Tampa, FL 33612-3805, USA

*Author to whom correspondence should be addressed. Tel: 813-974-6629; e-mail: tbernard@health.usf.edu

Submitted 13 November 2016; revised 28 March 2017; editiorial decision 28 March 2017; revised version accepted 11 May 2017.

Abstract

Objectives: There are times when it is not practical to assess heat stress using environmental metrics and metabolic rate, and heat strain may provide an alternative approach. Heat strain indicators have been used for decades as tools for monitoring physiological responses to work in hot environments. Common indicators of heat strain are body core temperature (assessed here as rectal temperatureT_{re}), heart rate (HR), and average skin temperature (T_{sk}). Data collected from progressive heat stress trials were used to (1) demonstrate if physiological heat strain indicators (PHSIs) at the upper limit of Sustainable heat stress were below generally accepted limits; (2) suggest values for PHSIs that demonstrate a Sustainable level of heat stress; (3) suggest alternative PHSIs; and (4) determine if metabolic rate was an effect modifier.

Methods: Two previous progressive heat stress studies included 176 trials with 352 pairs of Sustainable and Unsustainable exposures over a range of relative humidities and metabolic rates using 29 participants. To assess the discrimination ability of PHSIs, conditional logistic regression and stepwise logistic regression were used to find the best combinations of predictors of Unsustainable exposures. The accuracy of the models was assessed using receiver operating characteristic curves.

Results: Current recommendations for physiological heat strain limits were associated with probabilities of Unsustainable greater than 0.5. Screening limits for Sustainable heat stress were T_{re} of 37.5°C, HR of 105 bpm, and T_{sk} of 35.8°C. T_{sk} alone resulted in an area under the curve of 0.85 and the combination of T_{sk} and HR (area under the curve = 0.88) performed the best. The adjustment for metabolic rate was statistically significant for physiological strain index or $\Delta T_{re\text{-}sk}$ as main predictors, but its effect modification was negligible and could be ignored.

Conclusions: Based on the receiver operating characteristic curve, PHSIs (T_{re}, HR, and T_{sk}) can accurately predict Unsustainable heat stress exposures. T_{sk} alone or in combination with HR has a high sensitivity, and makes better discriminations than the other PHSIs under relatively constant exposure (metabolic rate and environment) for an hour or so. Screening limits with high sensitivity, however, have low thresholds that limit utility. To the extent that the observed strain is low, there is good evidence that the exposure is Sustainable.

Keywords: heat strain; heat stress; Physiological responses; ROC; sustainable; unsustainable

Introduction

The assessment of heat stress begins with recognition followed by an exposure assessment, frequently using a wet-bulb globe temperature (WBGT) method (NIOSH, 2016; ACGIH, 2017). There are situations when making a traditional exposure assessment is not practical (e.g. maintenance tasks, usual work conditions, etc.) and to provide some evidence that the heat stress is well managed. Heat strain indicators have been used for decades as tools for monitoring physiological responses to work in hot working environments and providing limits to exposures (Brouha, 1960; NIOSH, 1972, 1986, 2016; Dinman *et al.*, 1974; Horvath, 1976; Fuller and Smith, 1981; Logan and Bernard, 1999). Rather than using physiological responses to limit an exposure, our paper considers their use to confirm that the exposures are Sustainable. Common indicators of heat strain are body core temperature, heart rate (HR), and skin temperature (T_{sk}).

Rectal temperature (T_{re}) is the usual surrogate for core temperature. At rest, the average adult core temperature is 37.0 ± 0.7. (Tanner, 1951; Cranston *et al.*, 1954; Sund-Levander *et al.*, 2002). The WHO scientific group on heat stress suggested 38°C as a limit on deep body temperature for prolonged daily exposures to heavy work and recognized that 39°C was safe under closely monitored conditions (WHO, 1969; NIOSH, 2016). The ACGIH® Threshold Limit Value® for heat stress and strain (ACGIH, 2017) suggested a limiting body core temperature of 38.5°C, which allows a margin to safely leave a heat stress exposure (Bernard and Kenney, 1994). Malchaire *et al.* (2001) examined the literature for a limiting core temperature and concluded that temperatures ≥39°C were likely to be associated with excessive heat strain. This premise was underpinned by Sawka *et al.* (1992) who found that cases of exhaustion rarely occurred when T_{re} was <38°C, and all observed heat exhaustion cases occurred before reaching 40°C.

HR is another index of heat strain. Ostchega *et al.* (2011) reported an average resting HR of 73 ± 3 bpm for adults aged between 20 and 59 years. Brouha (1960) observed that HR during work and recovery varies according to work load and ambient condition; he found a linear relation between HR increments and ambient temperature. Maxfield and Brouha (1963) reported that during environmental stress, the recovery of HR was prolonged with the increase in work load and increase of environmental temperature. To maintain a compensable level of heat stress, WHO reported an HR of 120 bpm for young, healthy men exposed to steady moderate work (from their Fig. 2) (WHO, 1969). Minard *et al.* (1971) demonstrated that daily average HRs above 120 would lead to a loss of aerobic work capacity for steel workers over a shift. Based on the results of their study, Kuhlemeier and Wood (1979) recommended a maximum HR for prolonged work at 125 bpm. Bernard and Kenney (1994) suggested HR thresholds around 125 bpm for exposures of 90 min. ACGIH® recommended discontinuing a heat stress exposure (Unsustainable heat stress) if the worker presents a sustained HR ≥180 bpm minus the individual's age in year (e.g. 180 − age) (ACGIH, 2017), based on a heat stress management practice in Australia.

T_{sk} can be used in industrial settings for long periods (Van Marken Lichtenbelt *et al.*, 2006; Smith *et al.*, 2010). Pandolf and Goldman (1978) recommended that if the difference between T_{re} and T_{sk} is <1°C, the exposure to heat should be stopped; and NIOSH (2016) repeats that recommendation. Assuming that a core temperature limit of 38.0°C is a target, a T_{sk} of 37°C would be a reasonable limit. Because of the reference to the difference between core and T_{sk}, we considered the difference as a single metric.

Moran *et al.* (1998) proposed the physiological strain index (PSI). It uses HR and T_{re} to represent both the cardiovascular and thermoregulatory systems and assumes that both contribute equally to the strain by assigning the same weight function to each. PSI = $5\,(T_{ret} - T_{re0})/(39.5 - T_{re0}) + 5\,(HR_t - HR_0)/(180 - HR_0)$. PSI evaluates heat strain on a common scale of 0 to 10, where 0 represents no strain and 10 represents strenuous (near maximal) physiological conditions. Buller *et al.* (2008) suggested a limit of 7.5, which was a little lower than the limiting heat strain allowed by their instutional review board (PSI = 8), to classify a person as at risk. Using the ACGIH limits of 38.5°C and 140 bpm (for age = 40), the PSI value is 6.1. Using WHO's limit of 38.0°C and HR of 120 bpm as Sustainable limits, the equivalent PSI is 5.1.

Exposure assessment is the usual approach to determine if a heat stress condition is Sustainable or Unsustainable (Garzon et al., 2016). In general, physiological heat strain indicators (PHSIs) (T_{re}, HR, and T_{sk}) will increase with the level of heat stress. The association of PHSIs and metabolic rate (M) is difficult to assess, and many investigators agree that core temperature is mainly determined by M below certain environmental temperatures (Lind, 1963, 1963; Lind *et al.*, 1970; Kuhlemeier and Wood, 1979; NIOSH, 2016). On the other hand, T_{sk} is largely independent of M and mainly associated with environmental temperatures (Wyndham *et al.*, 1965; Nielsen, 1966). The goal of the current study was to determine if PHSIs could accurately discriminate Sustainable from Unsustainable heat exposure for deciding whether to perform an exposure assessment. There were four objectives for undertaking this study: (1) demonstrate whether common PHSIs at the upper limit of Sustainable heat stress were below generally accepted limits; (2) suggest values for common PHSIs that demonstrate a Sustainable level of heat stress; (3) suggest alternative PHSIs; and (4) determine if metabolic rate was an important effect modifier for PHSIs. The models were then considered in the context of performance, that is, ability to distinguish between Sustainable and Unsustainable and how they might be used to screen for potential heat stress exposures.

Methods

For the current study, physiological data collected from progressive heat stress trials at University of South Florida (USF) (Bernard *et al.*, 2005, 2008) were used. The USF progressive heat stress studies were approved by the USF instutional review board. The study protocols, participant data, and case definitions were provided in Part 1.

T_{re}, HR, average T_{sk} [a weighted average of skin thermistors located on the arm, chest, thigh, and calf according to Ramanathan (1964)]. There were 176 trials for 29 participants wearing woven cotton clothing. For each trial, there were two pairs of Sustainable–Unsustainable observations. Because the participants were their own control, the observations were dependent.

Bernard *et al.* (2005) exposed the study participants to an M fixed approximately 160 W m^{-2} to approximate moderate work, at three levels of relative humidity. In the other study, Bernard *et al.* (2008) exposed the subjects at a relative humidity of 50% and levels of metabolic rate 115, 175, and 250 W m^{-2}, to approximate light, moderate, and heavy work. All of the participants were acclimatized by 2-h exposures over five successive days to dry heat (50°C and 20% relative humidity) at 160 W m^{-2} while wearing shorts and tee shirt.

Statistical analysis

Once the characteristics of the variables were assessed using Proc Univariate SAS 9.4 (SAS Institute Inc., 2013), a method to determine physiological limits between Sustainable and Unsustainable heat stress was explored. Proc Log (SAS Institute Inc., 2013) was used to fit conditional logistic regression models with PHSIs as predictors and Unsustainable and Sustainable conditions as the dichotomous outcome. Several models were built, beginning with unadjusted models, which later were fitted with the other PHSI predictors, added one by one until the best combination of predictors was achieved, and led to increase the predictability of the model.

Each of the principal indicators was expressed as a percent ratio over a nominal range from rest to highest acceptable value based on our judgment. The baseline values for T_{re} and HR in the ratios and the computation of PSI were 37°C and 75 bpm (average adult values). In addition, two other strain indicators were included. The ratios and additional indicators are described here and in Table 4.

$$RT_{re} = 100\left[(T_{re} - 37)/(39 - 37)\right]$$

$$RHR = 100\left[(HR - 75)/(180 - 75)\right]$$

$$RT_{sk} = 100\left[(T_{sk} - 35)/(37 - 35)\right]$$

$$\Delta T_{re\text{-}sk} = T_{re} - T_{sk}$$

$$PSI = 5(T_{re} - 37.0)/(39.5 - 37.0) + 5(HR - 75)/(180 - 75)$$

Unadjusted conditional logistic regression models were fitted with a single continuous predictor (RT_{re}, RHR, RT_{sk}, $\Delta T_{re\text{-}sk}$, PSI) to assess its association with the outcome; and statistical significance was accepted at $P \le 0.05$.

To see if the ability to predict an Unsustainable exposure improved with multiple variables, combinations of the three principal predictors were tested with multivariate multiple regression models. Stepwise logistic regression (SAS Institute Inc., 2013) helped to determine which second predictor (RT_{re}, RHR, RT_{sk}) was added into the model, the statistical significance was kept at $P \le 0.05$. Receiver operating characteristic curves (ROCs) were generated for the adjusted models, keeping a cut point of 0.95 for sensitivity. The order of the predictors was changed, that is, the second predictor was assessed as main predictor to determine which order improved the model. A covariate was maintained in the model only if it increased the association between the main predictor and the outcome increased by more than

10% or if it increased its ability to predict Unsustainable conditions [increasing its area under the curve (AUC)] by 0.05. M was added to the unadjusted and to the adjusted models, to assess its effect and to evaluate if its addition into the models increased their predictability.

Depending on the AUC, the conditional logistic regressions' results and the stepwise regression, a third predictor was added to the models to evaluate if it increased the accuracy of the model's prediction. Table 3 reports only those models with the best ability to predict true Unsustainable exposures.

Due to the crossover design, the observations within trials were dependent and treated that way. Each subject completed three different trials; with either three levels of humidity or three levels of metabolic rate. The three trials represent repeated measures because of the dependent observations within the trials and the repeated observations on the participants, a conditional logistic regression model fitted with the binary outcome and a continuous predictor was used.

Predictor model building

The conditional logistic regression (SAS Institute Inc., 2013) used to analyze dependent observations does not provide intercept (α) values required to build models at any fixed probability. To overcome this limitation, and to be able to discuss cut points for comparison purposes, a four-step process was followed:

1. The weights associated with each independent variable were determined from the conditional logistic equation based on the full data set of three conditions (compensable, critical, and uncompensable) in 176 progressive heat stress trials. In general, the physiological metric (Ψ) was $\Psi = \beta_1 x_1 + \beta_2 x_2$. For T_{re} and HR for instance, $\Psi = \beta_1 \mathrm{R}T_{re} + \beta_2$ RHR. In other words, this step provided a data-driven weighting of the physiological values based on the available data.
2. A conditional logistic regression using Ψ as predictor was then used to obtain the ORs for the association with the outcome variable (Sustainable versus Unsustainable).
3. A second database of only critical conditions was used to estimate a threshold value for Ψ via the predicted probability of Unsustainable. Increasing values of Ψ approximated a dose-response curve. The logistic regression was used based on the progressive count of cases divided by 176 as the dependent variable. The result was $\log[p/(1 - p)] = \alpha + \beta\,\Psi$.
4. As examples of threshold values, Ψ at $P = 0.05, 0.25, 0.50, 0.75$, and 0.95 were determined from $\log(p/[1 - p]) = \alpha + \beta\,\Psi$.

Testing the effects of metabolic rate

The association between M and the outcome was assessed using it as main predictor; after that M was fitted into multiple conditional logistic regression models with every one of the PHSIs and their combination. In those models where M was statistically significant, an interaction term was added to look for effect modification.

Assessing model performance

The accuracy of each unadjusted model to discriminate Unsustainable versus Sustainable was assessed using a ROC curve. The AUC is a measure of the overall ability to discriminate between Sustainable and Unsustainable. A SAS ODS statement was used along with a request to produce the ROC (SAS Institute Inc., 2013). Working from the premise that an indicator function was screening method, a cut point of 0.95 for the ROC sensitivity was found.

Because T_{sk} was the best main predictor of Unsustainable in the present study, its ROC curve was used as the comparison reference. T_{sk}'s ROC curve was contrasted with the other models' curves to determine if different PHSIs differed one from the other. The AUCs, their correspondent confidence intervals (CI), and the P-values from the contrast between T_{sk}'s AUC and the other PHSI's AUCs were obtained using R (R Development Core Team, 2015) and following the method described by Wu and Wang (2011).

Results

Descriptive statistics

The progressive heat stress studies included 176 trials, with 352 pairs of Sustainable and Unsustainable exposures three levels of relative humidities and three levels of metabolic rate using 29 participants. The characteristics of the study's volunteers were summarized in the companion paper.

Table 1 provides descriptive statistics (mean, standard deviation, and 5%–95% quantiles) of the principal PHSIs (T_{re}, HR, T_{sk}, and M).

Model development

Conditional logistic regression models were fitted with RT_{re}, RHR, and RT_{sk} as well as $\Delta T_{re\text{-}sk}$, PSI, and M. The association between the outcome and RT_{re}, RHR, and RT_{sk} as well as $\Delta T_{re\text{-}sk}$ and PSI was found statistically significant ($P < 0.001$). M alone was not significantly associated with the outcome ($P = 0.13$). The odds ratios (ORs) and CIs for these models are provided in Table 2.

Table 1. Summary statistics for rectal temperature, heart rate, skin temperature, and metabolic rate by case status.

	Rectal temperature (°C)			Heart rate (bpm)			Skin temperature (°C)			Metabolic rate (W)		
N	Mean	SD	Quantiles 5%–95%	Mean	SD	Quantiles 5%–95%	Mean	SD	Quantiles 5%–95%	Mean	SD	Quantiles 5%–95%
273 Sustainable	37.6	0.31	37.1–38.1	108	17	82–136	35.6	0.9	34.0–36.9	326	104	165–506
255 Unsustainable	37.9	0.33	37.4–38.5	130	20	100–164	36.8	0.8	35.5–38.1	331	104	170–506

Table 2. Results from the PHSI models, their correspondent AUCs, their specificities and cut points at a sensitivity of 0.95, and *P*-values from the contrast between RT_{sk}'s AUC against the other models' AUCs.

Metrics	OR (CI)	AUC (CI)	Specificity and cut point at sensitivity = 0.95	AUC comparison to RT_{sk} *P*-value
Unadjusted models				
RT_{re}	1.1	0.74	0.15	<0.0001
	1.08–1.12	0.70–0.78	37.4	
RHR	1.16	0.80	0.36	0.03
	1.13–1.19	0.77–0.84	24.8	
RT_{sk}	1.05	0.85	0.41	---
	1.04–1.06	0.82–0.89	28.5	
$\Delta T_{re\text{-}sk}$	0.22	0.77	0.28	<0.0001
	0.16–0.29	0.73–0.81	0.67	
PSI	5.47	0.81	0.36	0.06
	4.06–7.38	0.77–0.85	2.80	
Adjusted models				
RT_{sk}+RT_{re}	1.05	0.88	0.47	0.001
	1.04–1.06	0.85–0.91		
RT_{sk}+RHR	1.05	0.89	0.56	0.0003
	1.04–1.06	0.86–0.92		
RHR+RT_{re}	1.14	0.81	0.47	0.08
	1.12–1.17	0.85–0.91		
RT_{sk}+RHR+RT_{re}	1.05	0.89	0.53	<0.0001
	1.04–1.06	0.87–0.92		
Models adjusted for *M*[a]				
$\Delta T_{re\text{-}sk} + M$	0.13	0.90	0.52	<0.0001
	0.089–0.192			
PSI + *M*	9.35	0.81	0.34	0.06
	6.34–13.79			

[a]*M* was found to be only a confounder.

The third objective of this study was to determine if a combination of predictors would improve performance. Six conditional regression models were fitted with the PHSI predictors (three principal predictors, $\Delta T_{re\text{-}sk}$, PSI, three pairs of predictors, and one triplet predictor) to assess which was the best combination to improve the predictability of each main model. For the pairs, we selected the pair that increased the association between the main predictor and the outcome by more than 10% or increased the AUC by at least 0.05.

The combination of RT_{sk}, RHR, and RT_{re} was found by the stepwise logistic regression as the best model to predict Unsustainable cases (OR = 9.65, CI 5.73–16.26), with an AUC of 0.89. Table 3 shows the results from the best combination for each one of the physiological indices.

Table 3. Estimates from the conditional regression models and logistic regression models.

PHSI	Alternative physiological metrics
T_{re}	$RT_{re} = [(T_{re} - 37)/(39 - 37)]\ 100$ $\log[p/(1-p)] = -3.90 + 0.109\ RT_{re}$
HR	$RHR = [(HR - 75)/(180 - 75)]\ 100$ $\log[p/(1-p)] = -3.99 + 0.1005\ RHR$
T_{sk}	$RT_{sk} = [(T_{sk} - 35)/(37 - 35)]\ 100$ $\log[p/(1-p)] = -3.27 + 0.0528\ RT_{sk}$
$\Delta T_{re\text{-}sk}$	$\Delta T_{re\text{-}sk} = T_{re} - T_{sk}$ $\log[p/(1-p)] = -3.49 + 2.38\ \Delta T_{re\text{-}sk}$
PSI	$PSI = [5\ (T_{re} - 37.0)/(39.5 - 37.0)] + [5\ (HR - 75)/(180 - 75)]$ $\log[p/(1-p)] = -4.36 + 1.28\ PSI$
T_{sk} and T_{re}	$\Psi(RT_{sk}, RT_{re}) = RT_{sk} + (0.0971/0.0467)\ RT_{re} = RT_{sk} + 2.08\ RT_{re}$ $\log[p/(1-p)] = -4.70 + 0.0348\ \Psi$
T_{sk} and HR	$\Psi(RT_{sk}, RHR) = RT_{sk} + (0.1414/0.0436)\ RHR = RT_{sk} + 3.2\ RHR$ $\log[p/(1-p)] = -5.10 + 0.0309\ \Psi$
HR and T_{re}	$\Psi(RHR, RT_{re}) = RHR + (0.0229/0.1330)\ RT_{re} = RHR + 0.17\ RT_{re}$ $\log[p/(1-p)] = -\ 4.16 + 0.0906\ \Psi$
T_{sk} and HR and T_{re}	$\Psi(RT_{sk}, RHR, RT_{re}) = RT_{sk} + 0.119/0.0453\ RHR + 0.0422/0.0453\ RT_{re}$ $= RT_{sk} + 2.63\ RHR + 9.32\ RT_{re}$ $\log[p/(1-p)] = -\ 3.54 + 0.036\ \Psi$

Table 4. Distribution of scores for each of the PHSIs for only the critical conditions based on logistic regression model in Table 3.

	Probability of unsustainable				
PHSI	0.05	0.25	0.50	0.75	0.95
T_{re}	37.2	37.5	37.7	37.9	38.3
RT_{re}	9	26	36	46	63
HR	86	105	117	128	147
RHR	10	29	40	51	69
T_{sk}	35.1	35.8	36.2	36.7	37.4
RT_{sk}	6	41	62	83	118
$\Delta T_{re\text{-}sk}$	2.7	1.9	1.5	1.0	0.2
PSI	1.2	2.5	3.52	4.63	6.24
$\Psi(RT_{sk}, RT_{re})$	50	103	135	167	220
$\Psi(RT_{sk}, RHR)$	70	129	165	201	260
$\Psi(RHR, RT_{re})$	13	34	46	58	78
$\Psi(RT_{sk}, RHR, RT_{re})$	17	68	98	129	180

Table 3 summarizes the PHSI models. The first row summarizes the model, which includes the conditional logistic regression β's for the combination models (i.e. β_2/β_1, β_3/β_1, as appropriate for two or three metric models). The second row is the outcome of the logistic regression on the PHSI. Table 4 presents the thresholds for each of the PHSIs based on the logistic regression models provided in Table 3.

Models adjusted for metabolic rate

M was first fitted as main predictor in a conditional logistic model, its association with Unsustainable was not statistically significant. Then *M* was assessed for confounding in all the PHSI models and was found as a confounder for the models with $\Delta T_{re\text{-}sk}$ and PSI as main predictors, changing the ORs by more than 10%. With regard to effect modification, the effect of *M* was negligible in both models. The AUC for $\Delta T_{re\text{-}sk}$ with M increased 0.03; and there was no change in AUC for PSI.

Model performance

Table 2 shows the ORs with CIs from the conditional logistic regression models. The accuracy of the main models to predict Unsustainable was assessed by AUC in Table 2 with the 95% CI. For the purpose of discussion, ROC's sensitivity of 0.95 was chosen as optimal operating point (Gallop, 2001) to reliably determine if an exposure was Sustainable.

Table 2 provides the *P*-values from the contrast between the PHSI models' AUC against the T_{sk}'s AUC. Fig. 1 contrasts the RT_{sk} ROC and the model fitted with RT_{sk} and RHR as predictors. There was approximately 0.04 difference between the AUCs.

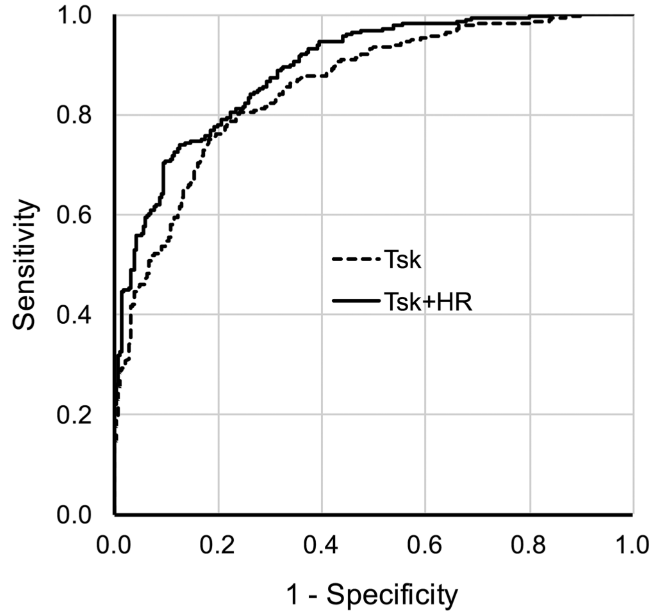

Figure 1. Contrast between ROCs for T_{sk} (AUC = 0.85) and T_{sk} + HR (AUC = 0.89).

Discussion

The clothing worn by the participants in this study was woven cotton (work clothes or coverall configurations). The mean metabolic rate found among the participants of the present study was 325 W (reported in Table 2 of Part 1), with data divided among three group averages of 115, 175, and 250 W m^{-2} and a range from 170 to 500 W (Bernard *et al.*, 2008). There was also a span of environments at 20, 50, and 70 relative humidity (Bernard *et al.*, 2005). This range of metabolic rates and environments spans the likely occupational exposures.

One purpose of the present study was to examine the distribution of PHSIs to see if they exceed commonly accepted physiological limits on heat strain. The results can also be used to assess if PHSIs or their combinations can accurately discriminate whether heat exposure is Sustainable for long periods. In this way, they might be a screening tool to determine if an exposure assessment is required. Inherent in the study design is the assumption that there is a steady-state physiological response. In more practical terms, we were looking at physiological metrics associated with a relatively constant exposure of an hour or more.

For the purposes of this paper, existing limits were compared to their probability of being associated with Unsustainable conditions; and the P = 0.25 for Unsustainable was used to describe a decision point for a screening method. As used in this study, probabilities of Unsustainable reflect indicator values along the dose–response curve (data at the critical condition only) rather than a probability from a random sample of observations. We used sensitivity and specificity to reflect our group of observations over compensable (all Sustainable), critical or transition (mix of Sustainable and Unsustainable), and uncompensable (all Unsustainable) exposures.

Comparison to current limits

The first objective of this study was to compare the distribution of PHSIs (specifically T_{re}, HR, and T_{sk}, $\Delta T_{re\text{-}sk}$, and PSI) to generally accepted limits. Table 4 provides a useful framework.

This often stated the goal of WBGT exposure limits to keep body core temperature below 38.0°C (NIOSH, 2016; ACGIH, 2017), with its origin in the WHO report (WHO, 1969). In this study, the higher predicted values of T_{re} in Table 4 were 38.0°C and 38.3°C. About 25% of the transitions from Sustainable to Unsustainable occurred above 38.0°C. This distribution with likely values greater than 38.0°C was consistent with Lind (1970) and Kuhlemeier *et al.* (1977). That is, the upper core temperature limit for Sustainable exposures is usually under 38.5°C, which was suggested by the ACGIH. We conclude that the 38°C limit was reasonable but should be thought more as a limit on the population average rather than an individual limit.

While several methods to use HR have been proposed, the only one that can be evaluated in this paper was a steady level. For this, the WHO (1969), Minard *et al.* (1971), Kuhlemeier and Wood (1979), and Bernard and Kenney (1994) recommend thresholds in the low 120s. From Table 4, the low 120s would put the predicted Unsustainable between 0.50 and 0.75 probabilities. Therefore, the upper limit guidelines have a high probability of Unsustainable.

There are no guidelines for T_{sk} alone. The present study collected average T_{sk}'s. At the higher probabilities of Unsustainable at 0.75 and 0.95 in Table 4, T_{sk} threshold would be 36.7°C and 37.4°C. At first glance, these might reasonably be at least 1°C below T_{re} recommended by Pandolf and Goldman (1978) with unacceptably high probabilities of Unsustainable. Instead of considering T_{sk} alone, $\Delta T_{re\text{-}sk}$ was selected to look more closely at the temperature gradient between core temperature and T_{sk}. The 1.0°C recommended by Pandolf and Goldman (1978) would result in a probability of Unsustainable of 0.75, which was too high for a Sustainable decision.

For PSI, Buller *et al.* (2008) suggested a limit of 7.5 to indicate risk at a high level of strain and the ACGIH physiological limits would translate to a value of 6.1. Both of these values represented a limit in a clearly Unsustainable exposure for most people.

In summary, current recommendations for physiological indicators were associated with probabilities of Unsustainable greater than 0.5.

Screening limits for sustainable heat stress

The second objective was to suggest values for PHSIs that demonstrate a Sustainable level of heat stress. We noted that at the PHSI cut points associated with sensitivity of 0.95 provided in Table 2 were about equivalent to a predicted probability for Unsustainable of 0.25. Thus, the predicted probability of 0.25 was used as the reference point for proposing screening limits.

The present investigation found that for the lower T_{re} for screening was 37.5°C. This value was in the range of resting core temperatures and represents a modest elevation over the average. This limits the utility of T_{re} as an indicator of Sustainable heat stress. By the same token, the lower HR in Table 4 is 105 bpm, which was substantially lower than the proposed limits discussed above. A candidate limit for T_{sk} was 35.8°C. This value gives a difference from the equivalent T_{re} (37.5°C) that is greater than 1°C, which suggested an adequate gradient for heat loss from the core to the skin. For $\Delta T_{re\text{-}sk} = 1.9$, the limit to predict a Sustainable level of heat stress agreed with Pandolf and Goldman (1978). For a reasonable threshold to limit Unsustainable decisions, the PSI was 2.5. This was substantially lower than the 5.1 that represents the more clearly Sustainable physiological strain.

In general, the individual physiological heat indicators were not practical predicators of Sustainable heat stress for potential use as a real-time administrative control. For long steady exposures to heat stress, $T_{re} < 37.5$, HR < 105, and $T_{sk} < 35.8$ were individually indicative of Sustainable heat stress. So for a protective screening decision, these have utility. That is, if any of the observed PHSIs is less than their threshold values, there is good reason to believe the exposure is Sustainable.

Models

To accomplish the third objective, we built four multivariate multiple conditional logistic regression models using PHSIs as predictors (see Table 2); and all of them demonstrated statistically significant association with the outcome of Unsustainable. The use of PHSIs such as T_{re}, T_{sk}, HR, and M or their combinations have been used by other investigators in different predictive equations to build their heat strain models (Pandolf *et al.*, 1986; Bernard and Kenney, 1994; Buller *et al.*, 2008; Cuddy *et al.*, 2013; Niedermann *et al.*, 2014; Richmond *et al.*, 2015).

At first glance, T_{re} was used both for the case definition of the outcome and as an independent variable in some of the models, which means that T_{re} should not be used in the models. In practice, the time course of T_{re} was used to demonstrate that the conditions were Sustainable or Unsustainable. The absolute value of T_{re} was only used as part of the case definition for Unsustainable in combination with the rate of increase.

Effect of metabolic rate

Metabolic rate is a measure of internal heat generation, and it is closely related with body core temperature (Nielsen, 1938; Saltin and Hermansen, 1966; Davies, 1979; NIOSH, 2016). The present study had the opportunity to assess PHSIs in a variety of combinations of light, moderate, and heavy metabolic rates. M was not statistically associated with Unsustainable. It was assessed for confounding and effect modification. M behaved as a confounder for $\Delta T_{re\text{-}sk}$ or PSI as main predictors. The confounding effect was related with the close association of M with body core temperature. Further, there was no meaningful contribution as an effect modifier. This finding was fortunate because M is very difficult to assess with any precision in field applications and thus not a practical consideration.

Comparisons among models

Care must be taken in comparing the ORs presented in Table 2. For continuous variables, the OR has an implicit denominator of 1 unit. For the percent ratios (RT_{re}, RHR, and RT_{sk}), the denominator is 1%. The denominator for PSI is 1 unit out of 10, and for $\Delta T_{re\text{-}sk}$, it is 1°C. To make the magnitudes roughly relatable to PSI, the percent ratios can be multiplied by 10 resulting in OR values around 11 compared to 5 for PSI. This process becomes too complex for multiple metrics. Again, this is offered here as rough illustration and should not be over-interpreted.

Tables 2 and 3 list nine models that might be used to predict the probability of Unsustainable heat stress exposure. The first three were single metric models based on the principal physiological indicators (T_{re}, HR, T_{sk}) expressed as percent ratios (RT_{re}, RHR, RT_{sk}); the next two were two-metric models ($\Delta T_{re\text{-}sk}$ and PSI) with a priori relationships between the two metrics; and four that were developed from the data based on the four combinations of the three principal metrics. Because the goal of this study is centered on the ability of PHSIs to distinguish between Sustainable and Unsustainable, we chose the AUC as the performance index by which to make the comparisons.

Among the principal physiological metrics, RT_{sk} clearly presented with the highest AUC at 0.85, which was significantly higher than RHR at 0.80 and RT_{re} at 0.74. As expected, the specificity at a sensitivity of 0.95 followed the same order at 0.41, 0.36, and 0.15, respectively. T_{sk} has not received much attention in the past as a potential indicator of heat strain because it does not correlate well with T_{re}. Both Bernard and Kenney (1994) and Richmond *et al*. (2015) found that insulated T_{sk} does not change much at lower levels of T_{re} but that it does have a monotonic increase as T_{re} increases. It might be its role in regulating heat exchange with the environment that allowed T_{sk} to play a dominant role in determining Unsustainable heat stress.

Because the difference between T_{re} and T_{sk} ($\Delta T_{re\text{-}sk}$) represented the potential for convective heat transfer from the core to the skin, it made a logical choice for assessment of heat strain (Pandolf and Goldman, 1978). The AUC for $\Delta T_{re\text{-}sk}$ was 0.77. For the purposes of comparison, the more general model that included both metrics was $\Psi(RT_{sk}, RT_{re})$ had an AUC of 0.88. Starting with the logit-p model in Table 4 and using P = 0.25, the relationship can be operationalized as

$$\log(0.25/0.75) = -1.10 = -4.70 + 0.038\Psi(RT_{sk},\ RT_{re})$$

$$\Psi(RT_{sk},\ RT_{re}) = 100\ [(T_{sk} - 35)/(37 - 35)] + 2.08 \times 100\ [(T_{re} - 37)/(39 - 37)]$$

$$3.6/0.038 = 95 = 100[(T_{sk} - 35)/(37 - 35)] + 2.08 \times 100[(T_{re} - 37)/(39 - 37)]$$

$$0.95 = [(T_{sk} - 35)/2] + 2.08[(T_{re} - 37)/2]$$

Values from the left side of the equation should be less than or equal to 0.95. For the P = 0.25 values of T_{sk} (35.8) and T_{re} (37.5), the combined metric value would be 0.92, which is similar to 0.95.

PSI was the other a priori two-metric indicator that was considered in this study. The AUC for PSI was 0.81 compared to 0.80 for RHR alone. That is, the contribution of T_{re} to the ability to discriminate between Sustainable and Unsustainable was relatively small. This observation was further supported by the same AUC for $\Psi(RHR, RT_{re})$.

Among the models with two predictors, $\Psi(RT_{sk}, RHR)$ had the best AUC at 0.89, with a specificity of 0.56. Because HR and T_{sk} are more accessible than core temperature, $\Psi(RT_{sk}, RHR)$ was a good candidate for assessing heat stress at the transition from Sustainable to Unsustainable. Several studies looking at the ability of HR and chest temperature to identify at-risk exposures (exposures that were greater than the transition heat stress levels of this paper) have demonstrated the utility of these two metrics together (Buller *et al.*, 2008; Cuddy *et al.*, 2013; Niedermann *et al.*, 2014). Operationalizing $\Psi(RT_{sk}, RHR)$ following the thought process described above,

$$129 = RT_{sk} + 3.2\ RHR = 100[(T_{sk} - 35)/(37 - 35)] + 3.2 \times 100\ [(HR - 75)/(180 - 75)]$$

$$1.29 = [(T_{sk} - 35)/2] + 3.2[(HR - 75)/105]$$

For the P = 0.25 values of T_{sk} (35.8) and HR (105), the combined metric value would be 1.31, which is similar to 1.29. Fig. 2 is a graphic representation of the $\Psi(RT_{sk}, RHR)$ limit using the data from this study.

The best combination of variables (T_{sk}, HR, and T_{re}), $\Psi(RT_{sk}, RHR, RT_{re})$, did not perform any better than $\Psi(RT_{sk}, RHR)$. Thus, the parsimonious model was RT_{sk}. The AUC could be improved from 0.85 to 0.89 by using the two-metric model that included T_{sk} and HR; that is, $\Psi(RT_{sk}, RHR)$. At a cut point associated with P = 0.25, the sensitivity for RT_{sk} was 0.93 with a specificity of 0.51. While $\Psi(RT_{sk}, RHR)$ had the same sensitivity of 0.93 and an improved specificity of 0.60, RT_{sk} was not just the best single predictor for Unsustainable cases, but its combination with RHR was a sensitive screening limit to identify Unsustainable exposures.

Limitations

There were three major limitations in this study. One was a data set designed to examine the transition from Sustainable to Unsustainable heat stress levels. For that reason, the conclusions were not generalizable to acute heat

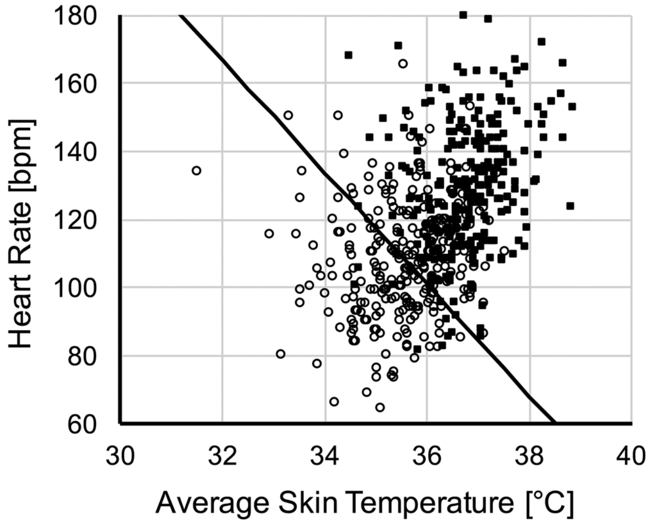

Figure 2. Illustration of the screening function using $\Psi(RT_{sk}, RHR)$. The open circles were Sustainable exposures and the closed squares were Unsustainable exposures.

stress and high, Unsustainable levels of heat stress. Another was the consideration of only single-layer woven clothing (e.g. work clothes or coveralls). The effects of nonwoven fabrics over a range of water vapor permeability and multiple layers were not considered. The third limitation was the practical consideration that the measurement is based on a relatively steady heat exposure for an hour.

Another possible limitation of this study as that the data obtained in both USF studies were collected on laboratory trials, under controlled conditions with acclimatized participants which were not similar than those present in real work settings, and therefore generalization could be affected. Nonetheless, this probable lack of generalization could have been attenuated by the fact that the study volunteers were exposed to a large range of metabolic rates (170 to 500 W) and environmental conditions (large range of humidity from 20% to 70% relative humidity).

Conclusions

The present study found that PHSIs (T_{re}, HR, and T_{sk}) can accurately predict Unsustainable heat stress exposures. T_{sk} alone or the combination of skin temperature and HR had the highest AUCs. In the context of the four research objectives:

1. The generally recommended values for the principal indicators (T_{re}, HR, and T_{sk}) and the two *a priori* indicators ($\Delta T_{re\text{-}sk}$ and PSI) were associated with probabilities of Unsustainable greater than 0.5, which meant that they were not sensitive to Sustainable exposures.
2. When the values of the principal and a priori indicators associated with a predicted probability of 0.25 were considered, they were generally much lower than the generally accepted physiological limits. The value of using any one of these individual indicators is that they act as a screening tool to decide if an exposure assessment is needed.
3. The most parsimonious model was RT_{sk} with an equivalent screening value of 35.8°C, which had a sensitivity of 0.93 and specificity of 0.51. The AUC could be improved from 0.85 to 0.89 by using $\Psi(RT_{sk}, RHR)$: $1.29 \le [(T_{sk} - 35)/2] + 3.2\ [(HR - 75)/105]$. The sensitivity and specificity of this two-metric indicator were 0.93 and 0.60, respectively.
4. Metabolic rate was found to be a confounder, but not an effect modifier, for $\Delta T_{re\text{-}sk}$ and PSI. Metabolic rate was not statistically significant with any other PHSI.

The results of this study suggested that PHSIs might be an intermediate step between recognition and exposure assessment. We believe that the methods in this paper of assessing the ability to discriminate are relevant to assessing the physiological indicators for other case definitions such as real-time monitoring to stop exposures.

Funding

Funding for this project was provided by the Republic of Ecuador and CDC/NIOSH [R01-OH03983 and T42-OH008438].

Acknowledgements

The data used in this study were collected under CDC/NIOSH R01-OH03983. Dr. Garzón was supported by the Republic of Ecuador through the Ecuadorian Institute of Human Talent Development. The authors recognize and thank the many laboratory assistants and trial participants who made this study possible. We would like to especially recognize Victor Caravello along with Mathew D. Dooris, Brian E. Grace, Patrick L. Rodriguez, and Brianna Clendenin Patton. Caravello was supported by the US Air Force, Dooris by the US Coast Guard, and Grace, Rodriguez and Patton were supported by CDC/NIOSH T42-OH008438.

Declaration

The authors declare no conflict of interest relating to the material presented in this article. One of the authors (TEB) has acted as an expert witness for both private companies and OSHA in litigation concerning heat stress exposures and may in future serve as an expert witness in court proceedings related to heat stress.

The contents, including any opinions and/or conclusions expressed, are solely those of the authors.

References

ACGIH. (2017) Heat stress, TLVs and BEIs: threshold limit values for chemical substances and physical agents & biological exposure indices. Cincinnati, OH: ACGIH.

Bernard TE, Caravello V, Schwartz SW *et al*. (2008) WBGT clothing adjustment factors for four clothing ensembles and the effects of metabolic demands. *J Occup Environ Hyg*; **5**:1–5; quiz d21–3.

Bernard TE, Kenney WL. (1994) Rationale for a personal monitor for heat strain. *Am Ind Hyg Assoc J*; **55**:505–14.

Bernard TE, Luecke CL, Schwartz SW *et al*. (2005) WBGT clothing adjustments for four clothing ensembles under three relative humidity levels. *J Occup Environ Hyg*; **2**:251–6.

Brouha L. (1960) *Physiology in industry. Evaluation of industrial stresses by the physiological reactions of the worker*. Oxford: Pergamon Press.

Buller MJ, Latzka WA, Yokota M *et al*. (2008) A real-time heat strain risk classifier using heart rate and skin temperature. *Physiol Meas*; **29**:N79–85.

Cranston WI, Gerbrandy J, Snell ES. (1954) Oral, rectal and oesophageal temperatures and some factors affecting them in man. *J Physiol*; **126**:347–58.

Cuddy JS, Buller M, Hailes WS *et al*. (2013) Skin temperature and heart rate can be used to estimate physiological strain during exercise in the heat in a cohort of fit and unfit males. *Mil Med*; **178**:e841–7.

Davies CT. (1979) Thermoregulation during exercise in relation to sex and age. *Eur J Appl Physiol Occup Physiol*; **42**:71–9.

Dinman BD, Stephenson RR, Horvath SM *et al*. (1974) Work in hot environments. I. Field studies of work load, thermal stress and physiologic response. *J Occup Med*; **16**:785–91.

Fuller FH, Smith PE Jr. (1981) Evaluation of heat stress in a hot workshop by physiological measurements. *Am Ind Hyg Assoc J*; **42**:32–7.

Gallop RJ. (2001) Determination and Interpretation of the OOP for ROC's with PROC LOGISTIC. *Proc NESUG*; 777–82.

Garzon XP, Wu Y, Ashley CD, Bernard TE. (2016) Ability of WBGT exposure limits to discriminate between sustainable and unsustainable heat stress exposures.

Horvath S. (1976) *Heat stress studies in aluminum reduction plants. Standards for occupational exposures to hot environments* (HEW Pub. No.[NIOSH] 76–100), SM Horvath (ed.) Cincinnati, OH: National Institute for Occupational Safety and Health.

Kuhlemeier KV, Miller JM, Dukes-Dobos FN *et al*. (1977) Determinants of the prescriptive zone of industrial workers. *J Appl Physiol Respir Environ Exerc Physiol*; **43**:347–51.

Kuhlemeier KV, Wood TB. (1979) Laboratory evaluation of permissible exposure limits for men in hot environments. *Am Ind Hyg Assoc J*; **40**:1097–103.

Lind AR. (1963) A physiological criterion for setting thermal environmental limits for everyday work. *J Appl Physiol*; **18**:51–6.

Lind AR. (1963) Physiological effects of continuous or intermittent work in the heat. *J Appl Physiol*; **18**:57–60.

Lind AR. (1970) Effect of individual variation on upper limit of prescriptive zone of climates. *J Appl Physiol*; **28**:57–62.

Lind AR, Humphreys PW, Collins KJ *et al*. (1970) Influence of age and daily duration of exposure on responses of men to work in heat. *J Appl Physiol*; **28**:50–6.

Logan PW, Bernard TE. (1999) Heat stress and strain in an aluminum smelter. *Am Ind Hyg Assoc J*; **60**:659–65.

Malchaire J, Piette A, Kampmann B *et al*. (2001) Development and validation of the predicted heat strain model. *Ann Occup Hyg*; **45**:123–35.

Maxfield ME, Brouha L. (1963) Validity of heart rate as an indicator of cardiac strain. *J Appl Physiol*; **18**:1099–104.

Minard D, Goldsmith R, Farrier PH Jr *et al*. (1971) Physiological evaluation of industrial heat stress. *Am Ind Hyg Assoc J*; **32**:17–28.

Moran DS, Shitzer A, Pandolf KB. (1998) A physiological strain index to evaluate heat stress. *Am J Physiol*; **275**(1 Pt 2):R129–34.

Niedermann R, Wyss E, Annaheim S *et al*. (2014) Prediction of human core body temperature using non-invasive measurement methods. *Int J Biometeorol*; **58**:7–15.

Nielsen B. (1966) Regulation of body temperature and heat dissipation at different levels of energy and heat production in man. *Acta Physiol Scand*; **68**:215–27.

Nielsen M. (1938) Die Regulation der Körpertemperatur bei Muskelarbeit1. *Skand Archiv Physiol*; **79**:193–230.

NIOSH. (1972). *Criteria for a recommended standard: occupational exposure to hot environments*. Washington D.C.: U.S. Department of Health and Human Services.

NIOSH. (1986). *Occupational exposure to hot environments*. Washington D.C.: U.S. Department of Health and Human Services.

NIOSH. (2016). *Criteria for a recommended standard: occupational exposure to heath and hot environments*. By

Jacklitsch B, Williams WJ, Musolin K, Coca A, Kim J-H, Turner N. Cincinnati, OH: Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, DHHS (NIOSH). Publication 2016-106.

Ostchega Y, Porter KS, Hughes J, Dillon CF, Nwankwo T. (2011) Resting pulse rate reference data for children, adolescents, and adults: United States, 1999–2008. *Natl Health Stat Report*; 1–16.

Pandolf KB, Goldman RF. (1978) *The convergence of skin and rectal temperatures as a criterion for heat tolerance. Aviat Space Environ Med*; **49**:1095–101.

Pandolf KB, Stroschein LA, Drolet LL *et al*. (1986) Prediction modeling of physiological responses and human performance in the heat. *Comput Biol Med*; **16**:319–29.

R Development Core Team. (2015) *R: A language and environment for statistical computing*. Vienna, Austria: *R Foundation for Statistical Computing*.

Ramanathan NL. (1964) A new weighting system for mean surface temperature of the human body. *J Appl Physiol*; **19**:531–3.

Richmond VL, Davey S, Griggs K, *et al*. (2015) Prediction of core body temperature from multiple variables. *Ann Occup Hyg*; **59**:1168–78.

Saltin B, Hermansen L. (1966) Esophageal, rectal, and muscle temperature during exercise. *J Appl Physiol*; **21**:1757–62.

SAS Institute Inc. (2013) SAS 9.4. Cary, NC, USA.

Sawka MN, Young AJ, Latzka WA *et al*. (1992) Human tolerance to heat strain during exercise: influence of hydration. *J Appl Physiol (1985)*; **73**:368–75.

Smith AD, Crabtree DR, Bilzon JL *et al*. (2010) The validity of wireless iButtons and thermistors for human skin temperature measurement. *Physiol Meas*; **31**:95–114.

Sund-Levander M, Forsberg C, Wahren LK. (2002) Normal oral, rectal, tympanic and axillary body temperature in adult men and women: a systematic literature review. *Scand J Caring Sci*; **16**:122–28.

Tanner JM. (1951) The relationships between the frequency of the heart, oral temperature and rectal temperature in man at rest. *J Physiol*; **115**:391–409.

van Marken Lichtenbelt WD, Daanen HA, Wouters L *et al*. (2006) Evaluation of wireless determination of skin temperature using iButtons. *Physiol Behav*; **88**:489–97.

WHO. (1969) *Health factors involved in working under conditions of heat stress*. Geneva, Switzerland: WHO Publications.

Wu Y, Wang X. (2011) Optimal weight in estimating and comparing areas under the receiver operating characteristic curve using longitudinal data. *Biom J*; **53**:764–78.

Wyndham CH, Strydom NB, Morrison JF *et al*. (1965) Criteria for physiological limits for work in heat. *J Appl Physiol*; **20**:37–45.