The reliability of local sweat rate measured via the ventilated capsule technique: effects of measurement region and level of heat strain

Maura McLean Rutherford
B.Sc., University of Ottawa

THESIS

In partial fulfillment of the requirements
Master of Science – Human Kinetics

Faculty of Heath Sciences, School of Human Kinetics
University of Ottawa, Canada

© Maura McLean Rutherford, Ottawa, Canada, 2020

ACKNOWLEDGEMENTS

I would like to express my endless gratitude to several individuals, without whom I would not have completed my thesis:

To Dr. Glen Kenny – thank you for the opportunity to complete my master's degree under your supervision. The support, encouragement and advice that you have offered me throughout my master's degree was invaluable to my learning. Your commitment to research, and to each of us, is admirable. The lessons I have learned throughout this degree, both academic and personally, will undoubtedly help me throughout my career and lifetime.

To my committee members, Dr. Heather Wright Beatty and Dr. Pascal Imbeault, thank you for the time and effort you have given to my thesis. Your feedback and guidance have helped me to be a better learner.

To the Human and Environmental Physiology Research Unit team, thank you for your support and friendship throughout the past two years. A special thank you to Dr. Sean Notley, Dr. Ashley Akerman, and Robert Meade for their guidance and knowledge, without which I would not be finishing this degree; and to Madison Schmidt, for being my partner in crime, and assistance in all things data.

I would like to thank all of the individuals who volunteered their time for my project. Without you, this thesis would not have been possible.

To my friends and family, thank you for encouragement throughout this journey. I am lucky to feel your support from provinces and countries away.

To Sean, who believed in me even when I didn't believe in myself and kept me grounded throughout this degree. Without you and your constant support this would have been impossible. I can't thank you enough.

Most importantly, thank you to my parents, Mark and Louise, and my brother Kewin. Thank you for your endless love and support, putting things into perspective when I needed it, and reminding me of what I am capable of. You encourage me to continue pushing myself and will always be the source of my motivation. You were vital to my thesis as you are with everything that I do, and for that, I thank you.

ABSTRACT

Ventilated capsules (i.e. hygrometry) are widely used to measure time dependent changes in local sweat rate. Despite this, understanding of the reliability (consistency) of local sweat rate is limited to the forearm during mild hyperthermia. Further, extensive regional heterogeneity in sweating may render some regions more reliable than others. Knowledge of reliability has important implications for experimental design, statistical analysis and interpretation, yet it is relatively unknown. The purpose of this study was to determine local sweat rate across various regions of the body and the reliability of these responses, during increasing levels of hyperthermia. On three separate instances, fourteen young men (age: 24 [SD 5] years) donned a whole-body water perfusion suit to raise and clamp esopogheal temperature at elicit low (+0.6°C), moderate (+1.2°C) and high (+1.8°C) levels of heat strain. Local sweat rate was measured at the forehead, chest, abdomen, bicep, forearm, hand, quadriceps, calf, and foot via ventilated capsules (3.8 cm^2). Absolute reliability was assessed using coefficient of variation (CV%) which quantifies the amount of error in a given measurement. Relative reliability was evaluated via the intraclass correlation coefficient (ICC); the consistency of an individual's rank within a group during repeated measurements. At low heat strain, most sites demonstrated acceptable relative (ICC ≥0.70), and moderate absolute reliability (CV <25%). At moderate-heat strain, the abdomen, hand, quadriceps, calf and foot had acceptable relative reliability while the forehead, abdomen, forearm, hand and quadriceps had moderate absolute reliability. At high-heat strain, relative reliability was acceptable at the abdomen, quadriceps, calf, foot and absolute reliability was moderate at the chest, abdomen, forearm, hand, quadriceps, calf and foot. Our findings indicate that reliability of

local sweat rate is dependent on both measurement site and level of hyperthermia. Researchers should consider this in their experimental design to increase the likelihood of detecting an effect of an intervention if one exists.

TABLE OF CONTENTS

PART ONE: THEORETICAL BACKGROUND

CHAPTER I: INTRODUCTION

1.1 Introduction

Humans strive to regulate body temperature within a narrow range. If the rate of heat loss mediated via increases in sweating and skin blood flow is not equal to that being gained via environmental exposure and/or exercise, body temperature (typically presented as mean body temperature, a weighted contribution of core and skin temperatures) will continue to rise. The hypothalamus then activates the thermoeffector responses responsible for heat loss in response to such an increase in body temperature, as sensed by central and peripheral thermoreceptors (Bligh, 2006; Gisolfi & Wenger, 1984; Kenny & Flouris, 2014).

Of the thermoeffector responses, humans rely on sweating (from eccrine sweat glands) during exposure to hot environments and/or exercise (Gagnon & Crandall, 2018); during which sweat rates can range from 0.5 – 1.5 $L \cdot h^{-1}$ and reach as high as 4 $L \cdot h^{-1}$ (Baker, Barnes, Anderson, Passe & Stofan, 2016). Sweating is not uniform across the body, with different sweat rates produced by various regions. The forehead and torso tend to have the highest sweat rates, while the distal extremities exhibit the lowest (Cotter, Patterson & Taylor, 1995; Smith & Havenith, 2011, 2012). The body temperature at which sweating is activated also differs across the body, with lower onset thresholds generally thought to occur in the thigh and back than the forearm (Smith, Alexander & Kenney, 2013). While some reports suggest that sweating onset occurs in a caudal to rostral sequence (Hertzman, Randall, Peiss, & Seckendorf, 1952; Park & Tamura, 1992; Rawson & Randall, 1961), others suggests that the pattern of response may be more complex (Cotter, Patterson & Taylor, 1995; Smith, Alexander & Kenney, 2013). Though sweat rate increases in relation to elevations in body temperature, skin sites such as

those of the feet and ankle demonstrate a plateau despite increasing heat loads. These findings suggest that regional heterogeneity exists in the maximal sweating response (Smith & Havenith, 2011).

The reliability, or consistency, of local sweating is relatively unknown despite widespread measurement of local sweat rate as a surrogate measure of heat dissipation. Reliability is important to establish for several reasons, including sample size calculations, interpreting results, and data analysis. Knowing how consistent or reliable a measure is and how much error can be expected provides important information needed to delineate if the observed response represents a true physiological response (i.e., activation of end organ – sweat gland) or is merely an artifact of measurement error. The assessment of local sweat rate (typically performed with using the ventilated capsule technique) has been widely assessed across different body regions for many decades, however only the reliability of the forearm has been determined (Kenefick, Cheuvront, Elliott, Ely & Sawka, 2012). While this has provided some insightful knowledge about the sweating response at measurement site commonly employed in many studies, it is not without its limitations. In their study, participants completed only one moderate-intensity exercise bout of walking for 30 minutes at 1.6 m/s and 5.0% grade. Consequently, they were only able to assess the reliability of the sweating response in context of a single exercise-induced heat load and therefore elevation in body temperature (increase in core temperature was limited to ~0.8°C). Given prior observations that regional heterogeneity exists in the sweating response in relation to changes in body temperature (Kondo et al. 1998), it is necessary to examine this response as a function of increasing elevations in heat load.

Moreover, given our limited knowledge of the pattern of response across the body, it is equally important that we assess this response across body regions.

1.2 Rationale

While the measurement of local sweat rate is widely used in thermoregulatory research to assess changes in thermoeffector activity, there remains a lack of knowledge on the reliability of this important measure. As outlined earlier, the reliability of local sweat rate technique as assessed by the ventilated capsule technique has been limited to assessment of the forearm (Kenefick et al. 2012). Yet, various reports demonstrate regional heterogeneity exists in the sweating response, which is in part attributed to differences in the number of sweat glands activated and/or sweat gland output (Kondo et al. 1998, 2001). It is well established that the increase in sweating is dependent upon the requirement for sweating as defined by elevations in environmental heat load (higher ambient temperatures increase the requirements for heat loss and therefore sweating), heat load generated by exercise (increases in metabolic rate augments the rate that heat must be dissipated) or their combination. While great strides have been made in advancing our understanding the human heat stress response as defined by our growing knowledge of the mechanisms governing the sweating response, there remains a significant gap in our understanding of how this response differs across the body. Understanding the regional pattern of response in sweating and the reliability of this response is especially critical given the importance of this commonly employed measure in human thermoregulatory studies.

1.3 Objectives

The primary objective of this study was to assess reliability of the sweating response across regions of the body using the ventilated capsule technique (hygrometry) as function of increases in heat load and therefore body temperature induced by whole-body heating. The specific study objectives were as follows:

1) To evaluate local sweat rate, onset threshold and thermosensitivity at the forehead, chest, abdomen, bicep, forearm, hand, quadriceps, calf and foot and determine the reliability of these responses, and;
2) To assess the degree to which the reliability of these sweating responses are dependent on increases in body temperature (low (0.6°C), moderate (1.2°C) and high (1.8°C) heat strain conditions) induced by whole-body heating.

1.4 Significance

Reliability statistics provide useful information to researchers to aid in sample size calculations, statistical analysis and interpretation of results. Understanding which regions are the most reliable at a given level of heat strain will help researchers select a site for assessment that is the most suitable to their experimental design. Further, by extending the knowledge of reliability to numerous regions across the body, there is a wider selection for researchers to choose from, as experimental design may exclude some regions from testing. This, in combination with information to distinguish between meaningful change and noise, will help researchers to be more confident that their effect is due to intervention, rather than noise.

CHAPTER II: REVIEW OF THE LITERATURE

2.1 Human Thermoregulation

Humans have developed a highly effective thermoregulatory system in order to maintain body temperature within the narrow range that is necessary for normal physiological functioning (Taylor, 2006). During heat stress, heat balance is achieved through adjustments in the heat loss responses to balance the rate of heat gained by the body with that lost to the surrounding environment. Heat loss is primarily achieved through the activation of skin blood flow and sweating, which facilitates increases in dry (via conduction, convection and radiation) and evaporative heat loss respectively (Gagge & Gonzalez, 1996). The rate of dry heat exchange is contingent on the temperature gradient between the environment and body surface (Kenny & Flouris, 2014). Increases in skin blood flow serve to increase skin temperature such that mean skin temperature exceeds ambient conditions, to favour dry heat loss (Gagge & Gonzalez, 1996). However, as ambient temperatures increase, the ability to lose heat via dry heat exchange is reduced as the skin-to-air temperature gradient narrows; when air temperature exceeds skin temperature, the body gains heat from the environment (Wendt, Van Loon & Lichtenbelt, 2007). Under these conditions, skin blood flow acts to bring heat to the skin's surface to be lost via the evaporation of sweat in order to offset dry heat gain (Benzinger, 1969; Gisolfi and Wenger, 1984; Hammel, 1968; Hardy, 1961).

Thermoregulation is mediated by a negative feedback loop with the pre-optic anterior hypothalamus acting as the 'thermostat' of the body (Benzinger, 1969; Hammel 1968; Hardy, 1961). Thermoafferent information is received from central and peripheral thermoreceptors in response to changes in body temperature, causing the pre-optic anterior hypothalamus to send efferent signals to initiate heat loss responses such as

increased skin blood flow and sweating (Hensel, 1981; Kenny & Flouris, 2014; Morrison & Nakamura, 2011). Thermoafferent activity is affected by both central (core) and peripheral (skin) thermoreceptors (Fusco, Hardy & Hammel, 1961; Kenny & Flouris, 2014). As illustrated in Figure 1, minor fluctuations in body temperature (recall, mean body temperature is a weighted contribution of core and skin temperatures) initially occur without the activation of heat loss responses. However, as the rate of heat gain increases, and mean body temperature rises more quickly, the activation of heat loss responses occurs which is defined as the onset threshold of sweating. Beyond the onset threshold, the change in effector response is proportional to changes in mean body temperature (Bligh, 2006). Thermosensitivity (Figure 1B) represents the change in the level of effector activity with increases in mean body temperature. However, despite further increases in mean body temperature, a plateau is achieved, there are no further increases in the heat loss responses (Figure 1C; Kenny & Flouris, 2014).

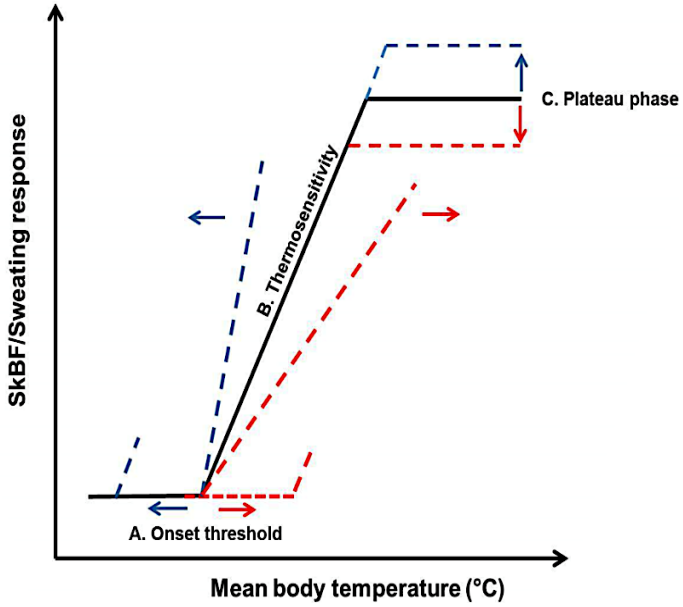


Figure 1. Schematic depicting the relationship between mean body temperature and the activation of thermoeffector responses of skin blood flow and sweating. Kenny & Flouris, 2014.

2.1.1. Evaporative Heat Loss

Evaporative heat loss (sweating) is one of the primary effector responses through which the body attempts to lose heat (Gagnon & Crandall, 2018). Sweating accounts for more than 80% of heat loss during exercise or environmental heat strain compared to only 25% at rest (Cain & McLellan, 1998). In response to mean body temperature surpassing the onset threshold (Figure 1A), the pre-optic anterior hypothalamus activates the 2-4 million eccrine sweat glands distributed across the body (Sato, Kang, Saga, & Sato, 1989) through cholinergic sympathetic nerves to increase sweat production and facilitate an increase in evaporative cooling (Shibasaki & Crandall, 2010). Humans have both apocrine and eccrine glands which differ in distribution, innervation and function. Apocrine glands are located in the underarm and genital regions, secreting fluid when stimulated by adrenergic sympathetic nerves (Sato et al. 1989). Only eccrine sweat glands are responsible for thermoregulatory sweating; however, they can also be stimulated by non-thermal factors such as exercise, psychological stress, and pain (Kondo et al. 2000; Machado-Moreira & Taylor, 2012a, 2012b; Sato et al. 1989; Van Beaumont, & Bullard, 1963). The sweating response varies in proportion to changes in mean body temperature; increases in sweating are initially associated with greater recruitment of sweat glands, followed by an increased sweat production from the recruited glands (Kondo et al. 2001). The recruitment of sweat glands is rapid, approaching maximal sweat gland recruitment within approximately 8 minutes of exercise or passive heat exposure, despite continued sweat gland output beyond this point (Kondo et al. 2001).

Prior to further discussion on the existing literature, it is first necessary to review the methodologies employed to measure sweating. Sweating was first assessed qualitatively, using visual detection (Buley, 1938; Kuno, 1938). This method was followed by photographic methods, in which photos of sweat glands were taken in succession to count the number of active glands, as well as estimate the output of these glands (Kenney & Fowler, 1988; Kuno, 1956). Measuring sweating has since become more quantitative, using gravimetry or hygrometry. Gravimetry is the measurement of collected sweat, used today primarily in the form of absorbent patches. A patch (or patches) are placed on the body for a short period of time (~5 min) to collect sweat as it is produced. The change in weight of the patch from pre- to post-application is used to quantify sweat production over that period of time. Given that saturation of the patch would lead to miscalculations in sweat rate, this technique can only be used for limited, single time points. Consequently, it is not possible to obtain an assessment of the time-dependent changes in sweat rate that typically occur. Further, this technique does not permit the assessment of changes in effector activity as represented by the evaluation of the onset threshold for the activation of sweating, and the thermosensitivity of the response. To facilitate these assessments, a temperature dependent profile of sweating is needed, which is only possible using the ventilated capsule technique, or hygrometry. Small capsules (~2-5 cm^2) adhered to the skin's surface have a known dry air (typically gas) moved through the capsule at a fixed flow, optimized for measuring local sweating. The expulsion of sweat increases the humidity (i.e., amount of moisture in the air) in the capsule environment and measuring the changes in the humidity of the effluent air relative to inflow corresponds to sweat rate (Brengelmann, McKeag & Rowell, 1975; Graichen,

Rascati & Gonzalez, 1982). Due to the sensitivity and precision of the ventilated capsule technique, they are the preferred method for measuring local sweat rate in mechanistic research (Taylor & Machado-Moreira, 2013).

2.1.2. Regional Variations in Sweating

There is regional heterogeneity in eccrine sweat gland density, so it follows that there would be regional variability in sweating. Regional heterogeneity in sweating is not only a function of the number of sweat glands in that area, but also the size and output of these glands (Kondo et al. 1998). It is established that regional variations in the size of sweat glands across the body exist; the larger the sweat gland the higher sweat rate and sensitivity of that gland (Sato & Sato, 1983). The rate of secretion from the sweat gland is determined by the sensitivity of the corresponding neurotransmitter (Kondo et al. 1998), with regional differences existing here as well. Regional variations in sweat gland density are partly determined by the intersegmental growth of an individual, as the number of sweat glands is pre-determined during gestation. As individuals grow and the surface area increases, density will decline. Low sweat gland densities are seen at the medial thigh and anterior leg, with high sweat gland densities on the chest, back and abdomen (Taylor & Machado-Moreira, 2013). An early report by Sato and Dobson (1970) showed that sweat gland density has the largest influence on a regional sweat rate, followed by the output and number of activated sweat glands in the area.

Knowledge of regional variations in sweating (as well as sweat gland density) between body segments to date is typically characterized by higher sweat rates on the torso and forehead, with lower sweat rates on the limbs (Cotter, Patterson & Taylor, 1995;

Smith & Havenith, 2011, 2012). Early studies showed that the activation of sweating occurs in a caudal to rostral pattern in young adults (Hertzman et al. 1952; Park & Tamura, 1992; Rawson & Randall, 1961). Later work has suggested otherwise; for example, Cotter, Patterson and Taylor (1995) reported that sweat onset at the torso preceded the head, but was not significantly different from the legs, upper torso or arms. Similarly, Smith, Alexander & Kenney (2013) demonstrated that the thigh and lower back had significantly lower onset thresholds than the forearm. More recent work has established the activation pattern is individualized and therefore random (albeit assessed only at four sites; Frei et al. 2019), demonstrating the need for more thorough investigation into the recruitment pattern of sweating.

With clear intersegmental differences in sweating, recent investigations have shifted to the assessment of intra-segmental variation. This is especially important given the fact that it is no longer acceptable to assume uniform distribution of sweating within any body segment (Machado-Moreria, Smith, van de Heuvel, Mekjavic & Taylor, 2008b). In this context, intra-segmental variations have been observed in torso and head regions. (Machado-Moreria et al. 2008b; Machado-Moreria, Wilmink, Meijer, Mekjavic & Taylor, 2008c). A study by Smith and Havenith (2011) undertook a comprehensive assessment of the intra-segmental sweating response by evaluating sweat production across the whole body. In their study, young endurance trained men performed two successive exercise bouts in the heat at increasing intensity (55 and 77% $\dot{V}O_{2max}$), during which time sweat production was measured using absorbent patches that covered approximately 83% of total body surface area. This showed that the highest sweat rates occurred in the torso and forehead, while the extremities exhibited the lowest sweat rates. Moreover, this

pattern of response remained intact with increases in intensity of exercise and therefore metabolic heat load. Consistent with the findings of Machado-Moreria and colleagues (2008b), they showed that the torso exhibited a medial to lateral decrease in sweating rate. Noteworthy, the sweat rate of the ankles and feet did not increase with exercise intensity, indicating maximal secretion at lower levels of heat stress for some regions, but not others. Despite a large inter-individual variation in absolute sweat rate a consistent pattern of regional variations was observed. However, due to the use of absorbent patches, regional variability of time-dependent profiles of sweating remains unknown.

2.2 Reliability Measures in Physiology

Reliability refers to the extent a variable or measure of interest is reproducible, or consistent. Reliability is assessed with repeated measurements, often completed over multiple trials. There are several variables of interest in reliability, but only those relevant for the purpose of this thesis will be discussed.

There are two major factors that affect reliability – physiological variation, and measurement/equipment error. Physiological variation (or commonly referred to as biological variation) stems from changes to a participant's physiological status or function. It is important for researchers to be aware of, and control for a myriad of factors that may affect an individual's physiologic function to the best of their ability in order to minimize biological variation. These can include physical activity, hydration status, caffeine intake, menstrual cycle, sleep schedule, physical and mental stress among others. Measurement or equipment error can be caused by different equipment being used during each trial or measurement, inaccurate or outdated calibration, and improper training.

Thus, it is important for researchers to be consistent in the equipment used and ensure it is properly calibrated. Ideally, the same researcher(s) who are properly trained, are operating the equipment for each trial.

Systematic differences or systematic bias is a tendency for measurements to change (increase or decrease) in a consistent direction when measured on repeat occasions. For example, if sweat rate was assessed on consecutive days, the repeated exposure to the heat would likely result in acclimation and would result in increasingly higher sweat rates each trial – a training effect. This can also be due to learning effects or lack of recovery between repeated trials (Atkinson & Nevill, 1998). Thus, researchers need to ensure there is adequate recovery between trials, and that trials are held at the same time of day to account for circadian rhythm (Atkinson & Nevill, 1998).These issues are addressed during the design phase of research; before determining the variability of a measurement, it is important to know if there are any systematic differences (changing baseline values) prior to undertaking a study.

Quantifying the contributions of error to variability in a measurement is done through absolute reliability. If an individual were to be passively heated for 30 minutes on three occasions and the sweat rate of the bicep was measured, how similar would the values to be to one another? The variation of these 3 sweat rates is referred to as ‘within-subject variation’, or ‘absolute reliability’, which is how much individual values differ when repeated (Atkinson & Nevill, 1998). The pooled standard deviation of the 3 local sweat rates represents the standard error in measurement, or the typical error (Hopkins, 2000). To facilitate comparison between measurement techniques (i.e., ventilated capsule vs absorbent patch) or different populations (i.e., young vs older adults), the typical error is

often presented as a percentage of the means, referred to as coefficient of variation (CV%; Hopkins, 2000). The lower the coefficient of variation for a given measure, the more reliable it is, as it represents a low amount of error in the measurement. Previous studies in the field of physiology have used the criteria of ≤10%, 10-25% and ≥25% coefficients of variation to be considered good, moderate and poor absolute reliability respectively (Iellamo et al, 1996; Tew et al, 2011). Of note, thresholds of 'good' reliability are not specific to any variable, nor are they standard across the field of physiology. What determines the reliability of a measure should be experiment specific and thresholds such as these are used for guides.

The minimum detectable change (MDC) can be calculated from the typical error of a measurement using a chosen confidence level, usually 95% (Beaton, 2000). The MDC is the absolute change that needs to be observed in a measurement, in order to have a given level of confidence that an effect is beyond the association error, and more likely due to an intervention. This is not to be confused with the minimally important difference (MID), or minimally clinically important difference (MCID) which is the minimum change needed to be observed in order to be considered physiologically or clinically important (Jaeschke, Singer, & Guyatt, 1989; Stratford, Binkley, Riddle, & Guyatt, 1998).

Another component of reliability is relative reliability or 'retest correlation', which is how well an individual maintains their 'rank' or position within a sample with repeated measurements (Atkinson & Nevill, 1998; Hopkins, 2000). For example, if there were 10 individuals performing the aforementioned experiment and participant 6 had the highest sweat rate at the bicep of the group, they would rank 1/10. If after repeated trials they remained at that rank, then there would be high relative reliability of that measure. The

interclass correlation coefficient (ICC) is used to measure relative reliability; the closer the ICC value is to 1, the higher the reliability. If participant 6 was ranked 1/10 for all three trials, the ICC value would be 1. There are different methods of calculating ICC, depending on the statistical design and intended interpretation and application of the results (Shrout & Fleiss, 1979). There are many available criteria as to what ICC values are considered reliable, but for the purposes of this thesis, ICC values ≥0.7 are considered suitable for research (Matheson, 2019). Similar to thresholds presented for 'good' absolute reliability, there are several thresholds used to interpret reliability that are neither field nor measurement specific.

Reliability has several uses in research. Firstly, assessment of reliability provides researchers with appropriate sample size calculations for their study. The typical error, or coefficient of variation is noise associated with a given change and knowing this gives researchers information needed to ensure the sample size is adequate to see real change outside of the noise in the measurement (Hopkins, 2000). Secondly, being aware of the typical error of measurement and using minimum detectable change in their analysis, researchers can be more confident that the change seen is due to intervention or factor being assessed (heat acclimation, hyperosmolality, aging).

2.2.1 Reliability of Sweating

The ventilated capsule is a common technique used to measure local sweat rates, due in part to the ease of use especially in context of measuring effector activity as defined by 1) the time dependent changes in relation to changes in thermal stimuli (e.g., exposure to hot environments and/or exercise; and 2) the assessment of the onset threshold and

thermosensitvity. The ventilated capsule technique itself is reliable as established by Kenefick and colleagues (2012), who determined the method has minimal contribution (a maximum of 3%) to variations in local sweat rate. However, while the technique has been proven to be reliable, little is known about the reliability of what this technique measures: local sweat rate, onset threshold and thermosensitivity. Recent work assessed the reliability of local forearm sweat rate and onset threshold during mild exercise induced hyperthermia (Kenefick et al. 2012). They report a moderate reliability (CV: 23.3%;) with a good reliability for the measurement of sweating onset (CV%: 9.6). Importantly, the authors assess local forearm sweating in the left and right arm, showing consistent responses. Consequently, their findings demonstrate that sweat rate and onset threshold measurements on the left and right sides of the body are interchangeable, which is important for study design when experimental set up may be limited or a control is needed (Kenefick et al. 2012).

Although this study was among the first to establish reliability of sweating at the forearm using the ventilated capsule technique, their assessment was limited to the forearm under a moderate heat stress conditions (i.e., core temperature increase of 0.8°C) induced by exercise. As outlined earlier, regional heterogeneity exists, and this response may vary as a function of elevations in heat load caused by exposure to hot environments and or exercise. In keeping with the latter, Kondo and colleagues (1998, 2001) showed that the level of heat stress can mediate the relative contribution of increases in sweat gland activation and sweat output to achieve the required increase in sweat rate. My project was conducted to explicate the impact of regional heterogeneity in sweating and heating level on reliability. Thus, this project was the first to establish

reliability of local sweat rate, onset threshold and thermosensitivity at various regions of the body, at multiple levels of heat strain.

PART TWO: METHODS AND RESULTS

NOTE: This is the pre-peer reviewed version of a submitted article under review. Document not to be distributed.

Myths and Methodologies: Regional variation in the reliability of sweat rate measured via the ventilated capsule technique during passive heating

Maura M. Rutherford*, Ashley P. Akerman*, Sean R. Notley, Robert D. Meade, Madison D. Schmidt and Glen P. Kenny

*Denotes joint first author

AUTHOR AFFILIATIONS:

Human and Environmental Physiology Research Unit, University of Ottawa, Ottawa, Ontario, Canada

ORCID:

MMR: 0000-0002-7916-6796

APA: 0000-0001-9194-2148

SRN: 0000-0002-5065-5000

RDM: 0000-0001-5922-4082

MDS: 0000-0001-6415-316X

GPK: 0000-0001-8683-6973

CORRESPONDING AUTHOR:

Glen P. Kenny,

125 University Private,

Room 367, Montpetit Hall,

Ottawa, ON,

Canada K1N 6N5

Email: gkenny@uottawa.ca

Phone: (+1) 613 562 5800 ext. 4282

Main body word count: ~8500

NEW FINDINGS

What is the central question of this study?

Does the measurement site and level of heat strain influence the reliability of local sweat rate measured via the ventilated capsule technique?

What is the main finding and its importance?

Reliability of sweat rate is dependent on site of measurement and magnitude of heat strain, such that sites meeting the pre-determined criteria of absolute or relative reliability generally declined from low- to high-heat strain. This information should be considered during both the design and interpretation of research utilising the ventilated capsule technique to assess eccrine sweating.

ABSTRACT

The ventilated capsule technique is widely used to measure time-dependent changes in sweating in humans. However, evaluations of its reliability (consistency) have been restricted to the forearm, despite extensive regional heterogeneity in the sweating response. Given the importance of such information for experimental design, statistical analysis, and interpretation, we determined the reliability of local sweat rate at nine sites during whole-body passive (resting) heating. On three separate occasions, a water-perfused suit was used to raise and clamp oesophageal temperature 0.6°C, 1.2°C, and 1.8°C above baseline in fourteen young men (age: 24 [SD 5] years), while sweat rate was measured at the forehead, chest, abdomen, bicep, forearm, hand, quadriceps, calf, and foot using ventilated capsules (3.8 cm^2). Absolute and relative reliability were determined via the coefficient of variation (CV) and intraclass correlation coefficient (ICC), respectively. At low-heat strain (0.6°C), almost all sites had acceptable relative reliability (ICC ≥0.70) and moderate absolute reliability (CV <25%). At moderate-heat strain (1.2°C), only the abdomen, hand, quadriceps, and foot had acceptable relative reliability, while the forehead, abdomen, forearm, hand and quadriceps had moderate absolute reliability. At high-heat strain (1.8°C), relative reliability was acceptable at the abdomen, quadriceps, calf and foot whereas the chest, abdomen, forearm, hand, quadriceps, calf and foot had moderate absolute reliability. Our findings indicate that the measurement site and level of heat strain impact the consistency of local sweat rate measured via the ventilated capsule technique, and thus, the likelihood of detecting an effect if one exists.

KEY WORDS: body temperature, heat stress, repeatability, reproducibility, sweat, thermoregulation.

LIST OF ABBREVIATIONS

CV: Coefficient of variation

ICC: Intraclass correlation coefficient

HHS: High heat strain

LHS: Low heat strain

LSR: Local sweat rate

MDC: Minimum detectable change

MHS: Moderate heat strain

NHS: No heat strain

SEM: Standard error of measurement

SMD: Smallest meaningful difference

TE: Typical error

INTRODUCTION

Humans rely heavily on eccrine sweat gland secretion to facilitate evaporative heat loss during heat stress (Gagnon & Crandall, 2018). Measuring sweat secretion is therefore of primary interest for researchers studying thermoregulation. While a variety of methods exist, the ventilated capsule technique is preferred for mechanistic research (Taylor & Machado-Moreira, 2013; Meade *et al.*, 2016), as it allows the user to; (i) obtain continuous measurement of sweat rate from multiple areas (~2-5 cm^2) on the body surface, (ii) identify the body temperature at which sweating begins (onset threshold), and (iii) quantify the increase in sweat rate per unit change in body temperature (thermosensitivity). Despite its widespread use, evaluations of its reliability (consistency) have been restricted to a single region (forearm) during low- (~0.3°C; Brengelmann *et al.*, 1994) to moderate-heat strain (~0.8°C; Kenefick *et al.*, 2012), or to assess the agreement between the ventilated capsule and the absorbent patch technique during exercise (Morris *et al.*, 2013). This is a major knowledge gap, as there is extensive regional heterogeneity in sweating (Taylor & Machado-Moreira, 2013) that may render some regions more reliable than others. Further, since the level of heat strain (indicated by change in core temperature) can modify sweat gland recruitment and output (Kondo *et al.*, 1998, 2001), the reliability of sweat rate at a given region may differ as a function of level of heat strain.

Given that such information is important for experimental design, statistical analysis, and data interpretation, we sought to evaluate the effects of measurement site and level of heat strain on the reliability of local sweating measured via the ventilated capsule technique. On three separate occasions, a water-perfused suit was used to raise

and clamp oesophageal temperature 0.6°C, 1.2°C, and 1.8°C above baseline, while sweat rate was recorded from nine sites. With this approach, we were able to quantify the reliability of the local sweating response at each site to incremental elevations in deep body thermoafferent drive that was both stable and matched across trials, while also quantifying the onset threshold and thermosensitivity of the response.

METHODS

Ethical approval

The experiment was approved by the University of Ottawa Health Sciences and Science Research Ethics Board (H-11-18-1186) and is in agreement with the latest revision of the *Declaration of Helsinki*, except for registration in a database. All participants provided written informed consent prior to participating.

Participants

The study was part of a larger project assessing the repeatability of commonly reported passive heat stress outcomes (Figure S1 in Supplemental Document; https://tinyurl.com/y2wfevwz). The minimum sample size (n=14) was calculated based on the expected precision (Shoukri *et al.*, 2004) of the reliability for baroreflex sensitivity, given this was the primary outcome for the broader project. Based on this sample size and a conservative estimate of relative reliability of sweat rate measured with an absorbent patch (intraclass correlation coefficient [ICC]: 0.90; Relf *et al.*, 2019) our design allowed for the estimation of the ICC to a precision (95% CI) of ~0.18 (see 'SampleSizeForReliabilityStudies.R' in the Supplementary material;

https://tinyurl.com/y4j9koaq). The study by Relf *et al.*, (2019) was selected, since ICC values for local sweat rate measured via the ventilated capsule technique were not provided in previous reports (Brengelmann *et al.*, 1994; Kenefick *et al.*, 2012; Morris *et al.*, 2013). Fourteen physically active men regularly performing ≥150 min/week moderate-to-vigorous activity, participated (Table 1). All were healthy (free of any cardiovascular, respiratory, autonomic or metabolic conditions), non-smokers and not taking any medications.

[INSERT TABLE 1 ABOUT HERE]

Experimental design

Participants completed one screening session (Table 1, and see Table S1 in Supplemental Document; https://tinyurl.com/y2wfevwz) and three identical experimental trials (separated by ~7 days) in a temperature-controlled laboratory (~25°C). Prior to each trial, participants were instructed to abstain from heavy exercise and alcohol for ≥24 hours, caffeine for at least four hours, and food consumption for at least two hours. Participants were also instructed to arrive at the laboratory adequately rested and having consumed 500 mL of water the night prior to, and the morning of the trial. Time of day, prior sleep, nutrition and physical activity habits were standardised within participants, and compliance checked verbally on arrival. Participants were reminded of their self-reported sleep duration, nutrition, and physical activity habits (for the 24 hours prior to the first trial) before the second and third trials.

Experimental trials

Upon arrival, each participant provided a urine sample to confirm euhydration (urine specific gravity: ≤1.025; Reichert TTS 400, Reichert, Depew, NY, USA; Kenefick & Cheuvront, 2012) and nude mass was recorded (model 2391; Detecto Scale Company, Webb City, MO, USA). Following procedures for another arm of the study (~30-45 minutes), which involved blood pressure measurements at rest and during ~5 minutes of squat-stand maneuvers (see Figure S1 in Supplemental Document for full details; https://tinyurl.com/y2wfevwz), participants were instrumented in the supine position wearing only shorts. Participants then donned a tube-lined water-perfusion suit covering the entire body except for the feet, head, and hands (Med-Eng, Ottawa, Canada), before resting for 10 minutes while baseline data were recorded (no-heat strain [**NHS**]). During this time, the suit was perfused (pump: GB series 81291, Micropump, WA, USA; drive: Model DJ605A, Micropump, WA, USA) with 34°C water (water bath: DC30-K20 Digitcal Control Bath, Thermo Scientific Haake, Germany). Water bath temperature was then increased to 49.5°C to raise and clamp oesophageal temperature at 0.6°C, 1.2°C, and 1.8°C above baseline. For brevity, these conditions will hereafter be referred to as low-, moderate- and high-heat strain [**LHS, MHS, HHS**], respectively. These qualitative statements reflect the relative heat strain in the current study, not generalisable or absolute levels of strain. Throughout heating, participants were covered to the neck with a plastic sheet, fleece blanket and a natural rubber mat.

At each level of heat strain, water-bath temperature was reduced to 43°C when oesophageal temperature was 0.1°C below the target temperature to elicit a plateau (~5 minutes). Data were then collected for 10 minutes before increasing water temperature

to 49.5°C to attain the next level of heat strain (Figure 1B). To standardise any non-thermal influences on sweating during each measurement period (Shibasaki *et al.*, 2006), breathing frequency was controlled (15 breaths/min) using a metronome and recorded via a respiratory belt (ADInstruments, Colorado Springs, CO, USA). After LHS and MHS, participants were given ~2 mL/kg and ~4 mL/kg, respectively of temperate water to limit body fluid losses. This quantity was chosen following pilot testing to limit body mass loss to <2%. Heating time to each level of heat strain was similar across trials (see Table 2 below).

[INSERT TABLE 2 ABOUT HERE]

EXPERIMENTAL MEASURES

Thermometry. Oesophageal temperature was measured using a Mon-a-therm General Purpose Temperature Probe (Mallinckrodt Medical, St Louis, MO, USA) inserted ~40 cm beyond the nostril and sampled at 1 kHz using an analog-to-digital converter (Powerlab, ADInstruments, Colorado Springs, CO) and stored for offline analysis (LabChart v8, ADInstruments). Skin temperature was measured at 1 minutes intervals using digital thermometers (iButton DS1921H-F5#, Maxim Integrated Products San Jose, CA, USA) affixed next to each ventilated capsule using hypoallergenic tape (Transpore; 3M, St. Paul, MN, USA) and used to derive mean skin temperature (forehead, 7%; forearm, 14%; hand, 5%; foot, 7%; calf, 13%; quadriceps, 19%; abdomen, 35%; Mitchell & Wyndham, 1969). Mean body temperature was estimated from the weighted sum of oesophageal (0.8) and mean skin temperatures (0.2; Hardy & DuBois, 1937).

Local sweat rates. Ventilated sweat capsules, each covering 3.8 cm^2 (Figure S2 in Supplemental Document; https://tinyurl.com/y2wfevwz), were attached to all sites with topical skin glue (Collodion HV, Mavidon Medical products, Lake Worth, FL, USA). The placement of each capsule was standardised using anatomical landmarks (Figure 1A). To minimise temperature and pressure artefacts between the water-perfused suit and sweat capsule (Frei *et al.*, 2019), perforated plastic domes (surface area of ~37 cm^2) were placed over each capsule. Dry-compressed air was supplied to the capsules at 0.75 L/min, while water content of the effluent air from each sweat capsule was measured with capacitance hygrometers (Model HMT333, Vaisala, Helsinki, Finland), which were calibrated per manufacturer specifications with standard salt solutions prior to experimentation (HMK15 Humidity Calibrator, Vaisala, Helsinki, Finland). Sweat rate was calculated every 5 s from the water content difference between effluent and influent air multiplied by the flow rate and normalized to the encapsulated skin surface area (mg/cm^2/min) and recorded with LabVIEW software (version 7.0; National Instruments, TX). To remove the influence of baseline variation, sweat rates at each site were normalised to the respective NHS values; however, all data are available via Supplementary Data Sheets Table S3 (https://tinyurl.com/y5d9x56a).

[INSERT FIGURE 1 ABOUT HERE]

Data analysis

Local sweat rates and oesophageal, mean skin, and mean body temperatures were expressed as minute averages, with the final 5 minutes of each 10-minute data

collection period being used for statistical analysis (Figure 1B). We also grouped local sweat rates into 3 regions (trunk: forehead, chest, abdomen; arm: bicep, forearm, hand; leg: quadriceps, calf, foot) to explore if averaging (unweighted) multiple sites would be beneficial to reliability. The onset threshold was determined as the mean body temperature at which a sustained increase in sweat rate >0.05 mg/cm^2/min occurred (Patterson *et al.*, 2004). Data are presented as the relative change in mean body temperature from NHS, but absolute data for onset threshold are available via Supplementary Data Sheets Table S5 (https://tinyurl.com/y5d9x56a). The thermosensitivity (slope) of the relationship between sweat rate and mean body temperature following the onset was determined using linear regression (Cheuvront *et al.*, 2009). The analyses for onset and thermosensitivity were conducted by a study author (MDS) blinded to the sites.

Statistical Analysis

Reliability and repeatability were assessed via estimates of systematic bias, absolute reliability, and relative reliability. The practical implications of these statistics on interpreting individual changes, sample size calculations, and the repeatability of the heating model are also presented. For detailed information related to these calculations, please see the Supplementary Document (https://tinyurl.com/y2wfevwz).

To describe any between-trial differences in thermometry measures, and the presence of systematic bias (i.e., the tendency for values to trend in the same direction due to a learning or training effect), data were analysed with a simple linear model at baseline and each level of heat strain, and the *p* value presented. Coefficient of variation (CV) is

presented as a measure of absolute reliability, or the magnitude of measurement error between repeated measures across trials. CV criteria of ≤10%, 10-25% and ≥25% were considered good, moderate, and poor, respectively (Iellamo *et al.*, 1996; Tew *et al.*, 2011). The standard error of measurement (SEM), or typical error (TE) from which CV was calculated is presented in the Supplementary Data Sheets Tables S2-S7 (https://tinyurl.com/y5d9x56a). Intraclass correlation coefficient (ICC) is presented as a measure of relative reliability of the rank order of the measurements in the form (ICC [3, 1]; Shrout & Fleiss, 1979), with ICC values ≥0.70 being considered suitable for research purposes (Nunnally, 1978; Matheson, 2019).

The minimum detectable change (MDC) was calculated as described by (Beaton, 2000), reflecting an estimate of the change that can be detected (beyond noise associated with error) between successive measurements at a given level of confidence (using the z-score consistent with a given confidence level; e.g., 1.96 for 95% confidence interval). MDC estimates are provided for different confidence levels (70%, 80%, 90%, 95% and 99%) to permit critical examination of the impact of measurement reliability on interpretation of data, and thus the ability for researchers to 'calibrate' their confidence in the data presented. To further emphasise this point, standardised effect sizes (Cohen's *d*) were calculated by expressing MDC estimates relative to the average between-subject SD, and sample size calculations were completed (assuming alpha: 0.05) to determine the required n for power ≥80% (beta: 0.20) for simple within- and between-subject designs. Thus, sample size calculations are provided to further illustrate this relation between measurement error and study design, and to highlight the mismatch in the current literature (measurements with large error being utilised to detect small differences,

or variable observed differences using these tools being interpreted as “true” effects). Notably, the average between-subject SD was chosen for ease of presentation (i.e., to calculate traditional Cohen’s *d*), but the variation associated with the change score (from NHS to LHS, MHS, and HHS) is provided in Supplementary Data Sheets Table S8 (https://tinyurl.com/y5d9x56a) for interested readers (and thus the ability to calculate Cohen’s *dz*). Lastly, to provide an indication of the replicability of the whole-body passive heating model on sweat rate at the group level, Cohen’s *d* effect sizes are presented for the change from NHS to each level of heat strain. Standardised effect sizes for comparisons between consecutive levels of heat strain (i.e., NHS-LHS, LHS-MHS, MHS-HHS) are presented in Supplementary Data Sheets Table S9 (https://tinyurl.com/y5d9x56a).

Absolute and relative reliability estimates were calculated using open source software (Hopkins, 2000; available via: https://www.sportsci.org/2015/ValidRely.htm) that was translated to a custom R (R Development Core Team, 2018) function for transparency and ease of calculation across multiple outcome measures. All other analyses were completed using packages in R. Further details are available via the Supplementary Document (https://tinyurl.com/y2wfevwz) and R scripts (https://tinyurl.com/y4j9koaq) related to the statistical analysis. Thermosensitivity calculations were generated using Prism 8 (GraphPad, CA, USA). Figures were produced using Prism 8 (GraphPad, CA, USA), and R (Power curves). All data are presented as mean (SD) or mean [95% CI], unless stated otherwise. For calculations of systematic differences, alpha was et at 0.05.

RESULTS

Local Sweat Rate

There were no systematic differences in local sweat rate (Table 3). Sweat rate across all sites ranged from 0.57 to 0.99 mg/cm^2/min at LHS (Figure 2) and increased to 0.71 to 1.29 mg/cm^2/min at MHS (Figure 3), and 0.73 to 1.35 mg/cm^2/min at HHS (Figure 4). The range of average absolute reliability (CV) between sites was largest at LHS (15 to 54%; Figure 2), but relatively consistent for MHS (Figure 3) and HHS (both ~15 to 30%; Figure 4). At LHS, average relative reliability (ICC) was ≥0.70 at all sites with the exception of the forehead, chest and bicep, whereas at MHS only the abdomen, hand, quadriceps, and foot were ≥0.70. With the exception of the hand, these sites as well as the calf were also the only sites with an ICC ≥0.70 at HHS (Figure 4).

[INSERT FIGURES 2-4 ABOUT HERE]

[INSERT TABLE 3 ABOUT HERE]

Regional Sweat Rate

Individual sites were grouped into trunk (abdomen, chest, forehead), arm (bicep, forearm, hand) and leg (quadriceps, calf, foot) regions to assess whether reliability was improved when a simple unweighted average was taken of the relevant sites. Each region exhibited moderate absolute reliability (14-22%), despite one site per region which did not meet this threshold as an individual site (Figure 2). At LHS, relative reliability was improved at the trunk (two individual sites did not have acceptable relative reliability) and

leg. At MHS, relative reliability improved in the trunk and leg while absolute reliability improved in all regions (Figure 3). Finally, at HHS, there were improvements in relative and absolute reliability for the trunk and leg (Figure 4).

Onset and Thermosensitivity

There were no systematic differences in the onset threshold or thermosensitvity (Table 4). The onset threshold ranged from 0.4 to 0.5°C for both individual sites and regional groupings (Figure 5A). Average absolute reliability of the onset threshold was poor at all sites and regions (CV <25%), and average relative reliability did not exceed (ICC) 0.70 at any site or region. Thermosensitivity ranged from 0.94 to 1.90 mg/cm^2/min/°C for individual sites, and 1.07-1.71 mg/cm^2/min/°C for regional groupings (Figure 5B). No sites or regions demonstrated average ICCs ≥0.70 or CV <25%.

[INSERT FIGURE 5 ABOUT HERE]

[INSERT TABLE 4 ABOUT HERE]

Minimum Detectable Change (MDC) and Sample Size Calculations

MDCs for all outcome variables, at each level of heating are presented in Figures 2-5. Based on these estimates, commonly reported effect sizes (~1.0 to 1.2; Inoue *et al.*, 1991; Shibasaki *et al.*, 2009; Kenefick *et al.*, 2012) and sample sizes (~10 to 20 for within- and between-group) in the literature, were associated with MDCs calculated at the 70 to 90% confidence levels, although this was dependent on the site/region and heating level.

If one wanted to ensure a change was "real" (i.e., beyond the noise of the measurement; at 95% confidence level) a standardised effect size (Cohen's *d*) of 1.1–2.6 (LHS), 1.3-2.7 (MHS), 1.3-2.5 (HHS) or 2.3–3.7 (onset, thermosensitvity) must be anticipated. Full sample size estimates and power curves are provided in the Supplementary Data Sheets Table S7 (https://tinyurl.com/y5d9x56a) and Supplementary material (https://tinyurl.com/y5xv96hc).

Replicability of the Model

The replicability of the whole-body heating on measures of eccrine sweating is presented in Figure 6 below. In lieu of a gold-standard approach to assess replicability, this assessment was based on the magnitude of the change from NHS to LHS, MHS, and HHS, and the associated 95% confidence interval in each trial. These data therefore encapsulate the range of effect sizes consistent with the data in each trial (at 95% confidence level; see Supplementary Data Sheets Table S8; https://tinyurl.com/y5d9x56a).

[INSERT FIGURE 6 ABOUT HERE]

As expected, there were consistent increases in sweating in response to the whole-body heating model. Point estimates of effect sizes at LHS spanned from large (Cohen's *d*: ~1.5) to very large (*d*: 5.3), but medium (*d*: ~0.7) to very large effects (*d*: 8.6) were consistent with our data (i.e., encapsulated in the 95% confidence interval). Point estimates of effect sizes during MHS and HHS were similar, ranging from large (*d*: 2.1)

to very large (*d*: 8.2), with large (*d*: 1.1) to very large effects (*d*: 12.6) within the 95% confidence intervals. The effect sizes were generally consistent for at least two out of three trials but appear to be site or region dependent. At each heating level there were select sites and a region that have point estimates and/or confidence intervals that were not consistent with other trials (LHS: forehead, abdomen, forearm, quadriceps, calf and arm; MHS: chest, abdomen, forearm, hand and trunk; HHS: forehead, chest, abdomen, forearm and trunk).

DISCUSSION AND PRACTICAL RECOMMENDATIONS

The following content serves two distinct purposes. Firstly, we highlight how the data obtained provide valuable insight into the reliability of sweat rate, onset and thermosensitivity, and how reliability changes with increasing levels of heat strain. This follows the framework of a 'traditional' discussion. Secondly, and relatedly, we provide a hypothetical example of how reliability data (such as those presented) may be used at different stages of the research process. The recommendations within this section combine specific data from the current study, and general comments on the importance of reliability in the research process.

Sweat rate, onset and thermosensitivity responses, and measurement reliability during whole-body passive heat stress

The current findings indicate that the reliability of sweat rate measured via the ventilated capsule technique is dependent upon both the site or region of measurement and level of heat strain. Average relative reliability (ICC) was acceptable for research

purposes (≥0.70) at most sites and all regions during low-heat strain, the abdomen, hand, quadriceps, foot and leg at moderate-heat strain, and the abdomen, quadriceps, calf, foot, trunk and leg at high-heat strain. Similarly, average absolute reliability (CV) was moderate (<25%) at most sites and all regions at low-, moderate-, and high heat-strain. These outcomes provide a valuable resource for experimental design, statistical analysis, and data interpretation for studies on eccrine sweating.

Brengelmann *et al.*, (1994) were the first to assess reproducibility of onset threshold during passive heating, however without dedicated reliability statistics comparison to current day assessments of reliability is problematic. The absolute reliability of sweating measured via the ventilated capsule technique has been previously assessed at the forearm during exercise-heat stress (Kenefick *et al.*, 2012), however the influence of level of heat strain or measurement site on both absolute and relative reliability remained uncertain. We therefore assessed the relative and absolute reliability at nine sites during three levels of heat strain. These results were then compared to previously employed thresholds (i.e., ICC ≥0.70 representative of acceptable relative reliability, and CV <25% representative of moderate absolute reliability; Table 5). With respect to relative reliability, the sites that met the research acceptable criteria of ICC values ≥0.70 were dependent on level of heat strain. At low-heat strain (Figure 2), all sites except for the forehead, chest and bicep were suitable, while at moderate-heat strain (Figure 3) the average relative reliability at the abdomen, hand, quadriceps and foot were ≥0.70. At high-heat strain, the abdomen, quadriceps, calf and foot were suitable for research purposes (Figure 4). No site met the ‘criterion’ value for good absolute reliability (coefficient of variation ≤10%; Iellamo *et al.*, 1996; Tew *et al.*, 2011), however most sites

exhibited moderate absolute reliability (10-25%) at low-heat strain (chest, abdomen, forearm, hand, quadriceps and calf; Figure 2). The forehead, abdomen, forearm, hand and quadriceps had moderate absolute reliability at moderate-heat strain (Figure 3), while the chest, abdomen, forearm, hand, quadriceps, calf and foot were moderately reliable at high-heat strain (Figure 4).

[INSERT TABLE 5 ABOUT HERE]

Averaging multiple sites into common anatomical regions (i.e., trunk, arm, leg) was explored to determine whether it provided any benefit to reliability. On several occasions a region was found to have met the respective threshold (i.e., ICC ≥0.70 and CV <25%) regardless of whether all three sites within that region did (i.e., at low-heat strain, the leg region exhibited moderate absolute reliability [14%] despite the foot having a CV of 53%). Additionally, if all three sites of a region met the criteria, the confidence intervals of the regional grouping appear to be narrower than that of the individual sites. The MDC estimates of the regional groupings were comparable to the individual site within that grouping that exhibited the lowest MDC estimate. Thus, averaging multiple sites provided slightly better (i.e., higher ICC and/or lower CV) and more precise (narrower 95% CI) estimates of reliability than that observed for the individual sites comprising that region. However, each region did contain a single site that exhibited acceptable relative and/or moderate reliability at each level of heat strain (abdomen [trunk], forearm/hand [arm], and quadricep [leg]). If resources and the experimental design permits, grouping multiple sites within a given anatomical region (i.e. trunk, arm, leg), would therefore account for intra-

regional differences in sweating without sacrificing reliability of the measurement. Nonetheless, regional reliability is dependent on region, level of heat strain and the reliability variable of interest.

Our findings also advance understanding by quantifying the reliability of the onset threshold and thermosensitivity of sweating. No sites or regions exhibited acceptable relative reliability (ICC ≥0.70) or moderate absolute reliability (CV <25%) for onset threshold (Figure 5A). Absolute and relative reliability of thermosensitivity was poor at all sites and regions (Figure 5B). These outcomes indicate that the reliability of sweating measured during the initial heating phase (i.e., transient conditions) is poorer than when under conditions of stable thermoafferent flow (i.e., during LHS, MHS, HHS oesophageal temperature clamps), perhaps due to the variation introduced when determining the slope of the relation between sweating and mean body temperature. As such, researchers interested in detecting within- or between-group differences in sweating may be more likely to detect a 'true effect' (if one exists) when those responses are assessed under conditions of matched and stable body temperatures rather than during thermal transients (i.e., local sweat rate rather than thermosensitivity).

Reliability estimates are seldom used to inform design, analysis, and interpretation despite their association (see below for further discussion). The reliability estimates allow one to determine the change (in any variable) that would need to be observed to have confidence that it was beyond the associated error (i.e., MDC). When determining the required sample size for a study (see Supplementary Material for Power curves; https://tinyurl.com/y5xv96hc), one can assess the likelihood of observing an effect that is likely to represent a real response (but this does not mean it is 'meaningful'). In the context

of the current study, our reliability estimates indicate that effect sizes with the highest levels of confidence (based on MDC) are likely to be attained at LHS and MHS (i.e., *d*=1.2 to 1.5; Inoue *et al.*, 1991; Shibasaki *et al.*, 2009; Kenefick *et al.*, 2012) but much less likely at HHS (i.e., *d*=1.3 to 2.5). Alternatively, these same data can be used to estimate the confidence one should have in their findings (i.e., 80% confidence when LSR differs by ~0.4 mg/cm^2/min at LHS or MHS).

For researchers wishing to determine the effects of heating on impairments or improvements to eccrine sweating, it is also necessary to employ an experimental protocol (model) that elicits consistent effects. The current results indicate that when assessed during whole-body passive heating, the range of effect sizes for sweat rate between trials within a particular site vary between 0.2-2.3, 0.4-4.5, and 0.1-3.0 Cohen's *d* at LHS, MHS and HHS, respectively (Figure 6). Thus, to ensure that the effect of the model on sweating is not only large, but consistently large, it is important to incorporate the level of heating and site of measurement into the experimental design using the data provided.

Study Conclusions and Practical Recommendations

Outcomes from this experiment indicate that there is no single 'most reliable site' to measure eccrine sweating via the ventilated capsule technique, with reliability being dependent upon the outcome measure (sweat rate, onset threshold, or thermosensitivity), site or region of measurement, and level of heat strain. Researchers should therefore consider these factors when designing research aimed at evaluating eccrine sweating to increase the likelihood of detecting an effect, should there be one. Specific practical

recommendations are presented below for using these reliability data to aid in research design and data interpretation. Next, a comprehensive worked example demonstrates specifically how one may consider reliability at different phases of the research process. Reliability, as determined by this study, is limited to healthy young men undergoing whole-body passive heating to three levels of heat strain. Thus, results may not be applicable to older adults, females, or individuals with chronic conditions. Further, as data were collected during passive heat stress using a water-perfused suit, results may differ during exercise or other variations of passive heating (water immersion, ambient exposure). Nonetheless, through these data one is able to better identify whether the anticipated effect of a study is beyond the noise of the measurement, what design characteristics are more favourable for identifying effects, and how to better interpret their (and others) findings with an appropriate degree of confidence (or uncertainty). Specifically;

A decision table has been made to assist researchers in choosing which site or region may be more reliable (based on a CV <25% and/or ICC ≥0.70) and appropriate for their design by incorporating heating level and sites available for measurement (Table 5). Each region (trunk, arm, or leg) has a site with acceptable relative and/or moderate absolute reliability. However, if the resources are available and the experimental design permits, grouping multiple sites in an anatomical region can provide a marginal benefit to reliability of local sweat rate while partly accounting for intra-regional differences in sweating.

For experiments involving high levels of heat strain (oesophageal temperatures >1.2°C above baseline), the effect sizes needed to have confidence in a true effect (exceeding measurement error) are larger than typically observed in the field. This implies

that for experiments where the anticipated effect is relatively small, lower levels of heat strain would be advisable.

Worked example for using reliability data in the research process

To improve research quality, knowledge of measurement reliability is critical. Unfortunately, these data are seldom considered beyond their absolute and relative reliability estimates (see (Atkinson & Nevill, 1998) for further discussion). However, by making use of these data one is able to better identify whether the anticipated effect of a study is beyond the noise of the measurement, what design characteristics are more favourable for identifying effects, and how to better interpret their (and others) findings with an appropriate degree of confidence (or uncertainty).

Factors associated with measurement error are associated with, but also independent of the statistical significance of any study. Unfortunately, most research in human physiology and many other fields (likely) suffer from low statistical power (Ioannidis, 2005; Button *et al.*, 2013; Fraley & Vazire, 2014; Abt *et al.*, 2020) and - counter to many researcher's intuitions - resulting in effect sizes that are variable, and wildly exaggerated when compared to the "true" underlying effect (Vul *et al.*, 2009; Gelman & Carlin, 2014; Loken & Gelman, 2017). When these observed effects are interpreted as an accurate representation of the "true" effect, we may unintentionally misrepresent the strength of evidence for a given phenomenon and lead future research down a path destined to fail.

Since no measurement tool/procedure has zero measurement error (i.e., perfect reliability), the data we observe does **not** represent the "true" effect, but rather an approximation of it. The difference between this approximation and the "true" value is dependent on the amount of measurement error present. This means that data must be

interpreted in the context of how much error accompanies it, and thus varying levels of uncertainty regarding the difference between the "true" and observed effects.

To guide researchers in how to incorporate reliability data in study design and data interpretation, we provide a simplistic example below. For clarity, the example will be referred to as the "hypothetical study", whereas reference to the data presented in the current manuscript will be referred to as the "current study".

Worked example: *Hypothetical study assessing sweat rate and heat acclimation*

Context

A researcher has developed a new heat acclimation protocol and wants to assess the impact it has on local sweat rate (a traditional marker of heat acclimation success). Local sweat rate (LSR) will therefore be assessed during a heat strain test before and after the acclimation protocol, exposing participants to a 1.0°C increase in oesophageal temperature via a whole-body water perfusion suit. To determine whether this project is feasible they must next identify whether: (i) a change is *likely* from a theoretical perspective, how *large* this change is likely to be, and if it represents a *meaningful difference*; and (ii) whether the tools available to them are *sufficiently sensitive (or reliable*) for these purposes.

Is the change likely and is the magnitude meaningful?

Typically, the most important information a researcher must first identify is the change in their outcome measure that represents the smallest meaningful difference

(SMD). Simply put, this value represents the smallest change that a researcher would not want to miss. In this example, our hypothetical researcher starts by considering what change in LSR would result in a meaningful impairment in whole-body sweat rate, increased body heat storage, or increased mean body temperature. However, these calculations are highly context specific, and hinge on several important assumptions, all of which may unintentionally introduce more error or uncertainty into the research process. Our hypothetical researcher therefore considers their experience with LSR measures and believes that a 0.20 mg/cm^2/min would be an appropriate approximation of a SMD. They justify this based on their knowledge of the intra-individual variation (i.e., SD) typically observed in LSR responses to heat stress (0.20 to 0.30 mg/cm^2/min; supported by data from the current study, see Supplementary Data Sheets Table S2; https://tinyurl.com/y5d9x56a). Their designated SMD would therefore represent between two-thirds to one SD in a similarly representative sample of participants.

Next, our hypothetical researcher examines the existing literature to determine how big of a change in LSR is likely in response to heat acclimation. Based on a meta-analysis of studies (Daanen *et al.*, 2018), our researcher feels that a 25% increase in sweat rate can be reasonably expected. To put this into the perspective for their own study, the researcher must make a few more choices about their research design. They decide that measuring LSR on the chest represents a common practice in the field, and logistically provides a convenient measurement site (based on the placement and access to the capsule under the water perfusion suit). The hypothetical researcher believes that the data presented in the current study is an appropriate representation of their anticipated LSR data. They therefore examine the MHS data of the current study (i.e.,

closest to the 1.0°C increase in oesophageal temperature in their intended heat strain test) and calculate that a 25% change (i.e., anticipated effect) would represent an absolute value of 0.31 mg/cm^2/min; (i.e., 1.23 mg/cm^2/min [average chest sweat rate at MHS] * 0.25; see Supplementary Data Sheets Table S2; https://tinyurl.com/y5d9x56a). Our hypothetical researcher is now at a point where their study of choice appears appropriate; the anticipated change (i.e., ~25%) exceeds what would be considered the SMD (i.e., 0.31 > 0.20 mg/cm^2/min, respectively). If their SMD was larger than the anticipated effect of the intervention, the study choice or specific design features may need reconsidering; even if they detect a statistically significant effect equal to anticipated effect, the real-world implications may be minimal and therefore unlikely to fully account for a physiological phenomenon (i.e., from a mechanistic perspective), or be used in any real-world context (i.e., from an applied perspective).

Is the effect likely to be observed with the tools available?

The next stage of the process is to determine whether our hypothetical researcher's research design and tools available are the most appropriate, or whether they are likely to detect a difference in LSR, if one exists.

As briefly highlighted above, observed measurements can be considered the combination of a "true" value and associated error, which could be owed to biological, technical, and random sources (Messick, 1987). A key process to determine approximately how much error is associated with a tool/procedure is to examine test-retest reliability. In particular, from estimates of absolute reliability (i.e., standard error of the measurement/typical error or coefficient of variation), a researcher can calculate the

minimum detectable change (MDC), or the change that is required to exceed the error associated with the tool/procedure (i.e., more reflective of a "true" difference, at a given confidence level). The calculation of MDC includes the z-score representing a given confidence level (e.g., 1.96 for 95% confidence interval; see '*Statistical Analysis' section* above for further details), meaning that a researcher can calibrate their degree of confidence or uncertainty in the data by calculating MDCs at different confidence levels. Thus, the MDC therefore provides a sign-post or reference point which can be used to aid research design, or interpret data that has already been collected (see '*Post-hoc use of reliability statistics'* below).

With this in mind, our hypothetical researcher considers the reliability estimates of the current study to be an accurate representative of their own test-retest reliability. They refer to the reliability estimates in Supplementary Data Sheets Table S2 (https://tinyurl.com/y5d9x56a) and the associated MDC at the 70% confidence level. The choice of the confidence level required, similar to other forking paths associated with research design (Gelman & Loken, 2013; Lakens *et al.*, 2018) is at the researcher's discretion. In instances where exploratory studies are being performed (for example), a researcher may be content with a lower confidence level, as the implications of the observed effect not closely representing the "true" effect are relatively minor. In circumstances where the absolute value of the measurement has important real-world implications, a higher confidence level would likely be required (e.g., measuring an individual's blood pressure to decide whether they start medication). When considering these implications, the hypothetical researcher is content that the MDC at the 70% confidence level is appropriate for their study.

In our current example, the hypothetical researcher now has estimates of the SMD, anticipated change, and MDC (0.20, 0.31, and 0.49 mg/cm^2/min, respectively). When considering these estimates, the feasibility of the experiment becomes much clearer. Namely, the researcher is reflecting on a question of **signal** (SMD or anticipated change) vs. **noise** (MDC). If, such as in this example, a MDC is larger than the anticipated change (or SMD) then the experiment is unlikely to be successful or worthwhile; even if a statistically significant effect is observed, this may merely reflect random variation (Loken & Gelman, 2017). Alternatively, a MDC that is lower than the anticipated effect (or SMD) is more likely to result in findings that represent true differences, and of a meaningful magnitude.

Our hypothetical researcher acknowledges that measuring LSR at the chest with the current experimental design therefore represents a problem; the MDC (0.49 mg/cm^2/min; Supplementary Data Sheets Table S2; https://tinyurl.com/y5d9x56a) exceeds the SMD and anticipated change (0.20 and 0.31 mg/cm^2/min; see above). In this situation, our hypothetical researcher must consider whether it is worth changing elements of their research design to account for these disparities and give their experiment the highest chance of detecting meaningful changes; i.e., increase the signal such that it is larger than the noise of the LSR measurement.

After consulting the practical recommendations presented in the current manuscript (see *'Study Conclusions and Practical Recommendations'* section above), our hypothetical researcher decides to measure LSR at the forearm instead of the chest. LSR at the forearm provides a more reliable estimate when individuals are exposed to an increase in oesophageal temperature of ~1.2°C. As above, if the current descriptive

statistics are assumed to provide an accurate representation of the researcher's sample of participants, the anticipated change (i.e., 25%; see above) now represents a change at the forearm of 0.26 mg/cm^2/min (1.03 mg/cm^2/min * 0.25; see Supplementary Data Sheets Table S2 (https://tinyurl.com/y5d9x56a). Due to lower measurement error at the chest, the MDC at the 70% confidence level now represents a 0.25 mg/cm^2/min change. Thus, our hypothetical researcher has now used the reliability estimates to adjust their design such that the anticipated signal is now beyond the noise of the measurement (i.e., anticipated effect > MDC), and proceeds to a power calculation for sample size determination. Considering that the SMD has not changed (0.20 mg/cm^2/min), this conclusion does not address the fact that the MDC is still larger than the SMD (i.e., 0.25 > 0.20 mg/cm^2/min), and this scenario will be covered below. However, it is worth highlighting that when a SMD for a study is greater than the MDC, we must critically evaluate whether the tools we are using are sufficiently sensitive to detect the effects we are hoping to see, particularly when designs become more complicated and we attempt to elucidate ever more complex physiological phenomenon.

Sample size calculation

A priori sample size determination, often via power calculations, is a necessary and important part of study design. However they are still rare in the field despite many valuable resources to aid in the process (Batterham & Atkinson, 2005; Lakens & Caldwell, 2019; Caldwell & Cheuvront, 2019). At the most basic level, a power calculation requires an estimate of the anticipated change (or difference), how much variability there is around this effect, and the error control rates (type 1 and type 2 error). To permit broad use,

power calculation software often uses standardised effect sizes rather than absolute values (i.e., dividing the effect by an associated variance component, such as a Cohen's *d or dz*). Researcher defined effect sizes are often derived from previous studies, or based on the history of sample sizes in the field, but as highlighted above, publication bias towards statistically significant findings and the tendency for underpowered study in this, and many other fields, means that this is not encouraged (Vul *et al.*, 2009; Button *et al.*, 2013). Pooled effect sizes from a meta-analysis (although this is prone to similar issues), and defining a SMD, means that researchers consider their topic more critically, particularly with respect to the smallest change they would not want to miss (i.e., SMD; Jones *et al.*, 2003; Batterham & Atkinson, 2005)

Our hypothetical researcher decides to complete a power calculation for the SMD (0.20 mg/cm^2/min) and the anticipated change (0.26 mg/cm^2/min) in LSR when individuals are exposed to whole-body passive heat stress before and after their heat acclimation protocol. Using the pooled between-subject standard deviation of forearm sweat rate at MHS (in lieu of a variance in the effect; i.e., 0.24 mg/cm^2/min, see Supplementary Data Sheets Table S2 (https://tinyurl.com/y5d9x56a), our hypothetical researcher calculates the standardised effect size (Cohen's *d*) for their SMD as 0.83 (i.e., 0.20 / 0.24 mg/cm^2/min), and for the anticipated change as 1.08 (i.e., 0.26 / 0.24 mg/cm^2/min). At this stage, they can again make use of the reliability estimates from the current study. Following the same procedure as above, the MDC at the 70% confidence level (0.25 mg/cm^2/min) represents a standardised effect size of 1.04 (*d*). This therefore represents a more conservative estimate than the standardised effect of the anticipated change (*d*: 1.08), but still a larger effect than the SMD (*d*: 0.83).

The researcher must therefore consider their options and the different possible forking paths in the research process. Powering their study to detect an effect equal to the MDC will permit a critical examination of the effect (if it is truly present), as they are considering measurement error and calibrating their confidence in the response after the fact. However, if the "true" effect is in fact smaller than the MDC, and in fact closer to the SMD, they may miss it and conclude that there is no difference despite the presence of a smaller, but still meaningful effect. Alternatively, powering their study to detect an effect equal to the SMD will require tools that can confidently detect these changes due to lower measurement error. In this sense, the researchers (rightly or wrongly) determine that they cannot access more reliable tools, and that the study they now have planned represents the best balance they are likely to encounter. Thus, they use a conservative estimate of the effect size (i.e., the MDC at 70% confidence level; *d*: 1.04, see above), and they proceed to the provided power curves (see Supplemental Material; https://tinyurl.com/y5xv96hc) to determine how many individuals are required with 80% power (beta: 0.2) and a critical alpha value of 0.05. In this instance, they identify that 10 participants are required for this within-subject design (simple t-test; see Figure 7 below).

[INSERT FIGURE 7 ABOUT HERE]

This hypothetical situation identifies the multiple forking paths when one is calculating sample sizes for a study. Often the process involves weighing the positives and negatives of different approaches *a priori*. However, in practice these complex considerations often fail to account for measurement reliability and instead assume that

our measurement tools provide perfect representations of “true” values; an assumption that is never true. Understanding this concept, and the potential ramifications for effects in the field to be exaggerated in their magnitude (Type M error) and the direction of the effect (or sign; Type S error; Gelman & Carlin, 2014) is critical if we anticipate that our studies will be replicated and stand the test of time.

The incorporation of reliability statistics into the power calculation procedure therefore permits a much broader understanding to our hypothetical researcher. With these tools they are able to identify an effect which represents a change beyond the noise of the measurement (at a given confidence level), and the knowledge of how much uncertainty they should have in those effects (point estimates and the associated confidence interval) if they are observed. Thus, considering measurement error and its inherent relation to study design, may ensure a more critical examination of the research process, and reduce the likelihood of wasted resources and missed opportunities.

Post-hoc use of reliability statistics

The same critical process is recommended after the fact, when data has already been collected, or a researcher is evaluating the available evidence of a phenomenon in the literature. Regardless of the *p*-value calculated, critical examination of the range of effects consistent with the data (i.e., point estimate and confidence intervals) provides more detailed insight into the evidence provided by studies. Relating observed effects to the anticipated measurement reliability, and meaningfulness of the absolute effect size will provide much more insight than mere reliance on a single *p*-value.

For instance, if an observed effect (and confidence interval) is statistically significant but the magnitude is below the MDC, a researcher should be cautious about over-interpreting their findings, or those of others; the MDCs (at different levels of confidence) therefore permit the researcher to calibrate the confidence in whether an effect is beyond the noise of the measurement or not, regardless of the statistical significance. On the other hand, if an observed effect is not statistically significant but the confidence interval encapsulates values which are above the MDC (and/or SMD), the phenomenon may still have promise, but the current findings may suffer from low statistical power. In this respect, over-reliance on dichotomised results (i.e., significant vs. not), rather than critical interpretation of data is likely a contributing factor to the poor reproducibility rate of scientific findings (Open Science Collaboration, 2015). These recommendations therefore add to growing calls to improve practices in the field (Caldwell *et al.*, 2020; Borg *et al.*, 2020) particularly with respect to acknowledging the role of measurement error (and reliability) to aid in robust research design, and critical evaluation of data, which may improve the inferences we make.

AUTHOR CONTRIBUTIONS

All authors contributed to conception and design of the work, data acquisition, analysis, or interpretation of the work, and contributed to drafting and critical revision of the manuscript.

COMPETING INTERESTS

The authors do not have any conflicts of interest to declare.

FUNDING INFORMATION

The work was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC; Discovery Grant no. RGPIN-2020-03891). M.M.R is supported by a Natural Sciences and Engineering Council of Canada Alexander Graham Bell Graduate Scholarship. A.P.A and S.R.N are supported by Postdoctoral Fellowships from the Human and Environmental Physiology Research Unit. R.D.M is supported by an Ontario Graduate Scholarship. G.P.K is supported by a University of Ottawa Research Chair Award.

ACKNOWLEDGEMENTS

The authors would like to acknowledge all participants, as well as Dr. Gregory McGarr and Mr. Mohamed Gemae for their assistance with data collection.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s). The data supporting the findings of this study, and the companion studies, is available from the supplementary material host (https://tinyurl.com/yys8rqpt).

REFERENCES

Abt G, Boreham C, Davison G, Jackson R, Nevill A, Wallace E & Williams M (2020). Power, precision, and sample size estimation in sport and exercise science research. *Journal of Sports Sciences*1–3.

Atkinson G & Nevill AM (1998). Statistical Methods For Assessing Measurement Error (Reliability) in Variables Relevant to Sports Medicine: *Sports Medicine* **26,** 217–238.

Batterham AM & Atkinson G (2005). How big does my sample need to be? A primer on the murky world of sample size estimation. *Physical Therapy in Sport* **6,** 153–163.

Beaton DE (2000). Understanding the Relevance of Measured Change Through Studies of Responsiveness: *Spine* **25,** 3192–3199.

Borg DN, Bon JJ, Sainani K, Baguley BJ, Tierney NJ & Drovandi C (2020). *Sharing Data and Code: A Comment on the Call for the Adoption of More Transparent Research Practices in Sport and Exercise Science*. SportRxiv. Available at: https://osf.io/ftdgj [Accessed August 26, 2020].

Brengelmann GL, Savage MV & Avery DH (1994). Reproducibility of core temperature threshold for sweating onset in humans. *Journal of Applied Physiology* **77,** 1671–1677.

Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ & Munafò MR (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci* **14,** 365–376.

Caldwell AR & Cheuvront SN (2019). Basic statistical considerations for physiology: The journal *Temperature* toolbox. *Temperature* **6,** 181–210.

Caldwell AR, Vigotsky AD, Tenan MS, Radel R, Mellor DT, Kreutzer A, Lahart IM, Mills JP & Boisgontier MP (2020). Moving Sport and Exercise Science Forward: A Call for the Adoption of More Transparent Research Practices. *Sports Med* **50,** 449–459.

Cheuvront SN, Bearden SE, Kenefick RW, Ely BR, DeGroot DW, Sawka MN & Montain SJ (2009). A simple and valid method to determine thermoregulatory sweating threshold and sensitivity. *Journal of Applied Physiology* **107,** 69–75.

Daanen HAM, Racinais S & Périard JD (2018). Heat Acclimation Decay and Re-Induction: A Systematic Review and Meta-Analysis. *Sports Med* **48,** 409–430.

Fraley RC & Vazire S (2014). The N-Pact Factor: Evaluating the Quality of Empirical Journals with Respect to Sample Size and Statistical Power ed. Ouzounis CA. *PLoS ONE* **9,** e109019.

Frei R, Notley SR, Taylor EA, Burdon CA, Ohnishi N & Taylor NAS (2019). Revisiting the dermatomal recruitment of, and pressure-dependent influences on, human eccrine sweating. *Journal of Thermal Biology* **82,** 52–62.

Gagnon D & Crandall CG (2018). Sweating as a heat loss thermoeffector. In *Handbook of Clinical Neurology*, pp. 211–232. Elsevier. Available at: https://linkinghub.elsevier.com/retrieve/pii/B9780444639127000138 [Accessed July 30, 2020].

Gelman A & Carlin J (2014). Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors. *Perspect Psychol Sci* **9,** 641–651.

Gelman A & Loken E (2013). The garden of forking paths: Why multiple comparisons can be a problem, even when there is no “fishing expedition” or “p-hacking” and the research hypothesis was posited ahead of time. *Department of Statistics, Columbia University*.

Hardy JD & DuBois EF (1937). Regulation of Heat Loss from the Human Body. *Proceedings of the National Academy of Sciences* **23,** 624–631.

Hopkins WG (2000). Measures of Reliability in Sports Medicine and Science: *Sports Medicine* **30,** 1–15.

Inoue Y, Nakao M, Araki T & Murakami H (1991). Regional differences in the sweating responses of older and younger men. *Journal of Applied Physiology* **71,** 2453–2459.

Ioannidis JPA (2005). Why Most Published Research Findings Are False. *PLoS Med* **2,** e124.

Jones SR, Carley S & Harrison M (2003). An introduction to power and sample size estimation. *Emergency Medicine Journal* **20,** 453–458.

Kenefick RW & Cheuvront SN (2012). Hydration for recreational sport and physical activity. *Nutrition Reviews* **70,** S137–S142.

Kenefick RW, Cheuvront SN, Elliott LD, Ely BR & Sawka MN (2012). Biological and analytical variation of the human sweating response: implications for study design and analysis. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* **302,** R252–R258.

Kondo N, Shibasaki M, Aoki K, Koga S, Inoue Y & Crandall CG (2001). Function of human eccrine sweat glands during dynamic exercise and passive heat stress. *Journal of Applied Physiology* **90,** 1877–1881.

Kondo, Takano, Aoki, Shibasaki, Tominaga & Inoue (1998). Regional differences in the effect of exercise intensity on thermoregulatory sweating and cutaneous vasodilation. *Acta Physiologica Scandinavica* **164,** 71–78.

Lakens D et al. (2018). Justify your alpha. *Nat Hum Behav* **2,** 168–171.

Lakens D & Caldwell AR (2019). *Simulation-Based Power-Analysis for Factorial ANOVA Designs*. PsyArXiv. Available at: https://osf.io/baxsf [Accessed August 26, 2020].

Iellamo F, Legramante M, Raimondi G, Castrucci F, Massaro M & Peruzzi G (1996). Evaluation of reproducibility of spontaneous baroreflex sensitivity at rest and during laboratory tests: *Journal of Hypertension* **14,** 1099–1104.

Loken E & Gelman A (2017). Measurement error and the replication crisis. *Science* **355,** 584–585.

Matheson GJ (2019). We need to talk about reliability: making better use of test-retest studies for study design and interpretation. *PeerJ* **7,** e6918.

Meade RD, Louie JC, Poirier MP, McGinn R, Fujii N & Kenny GP (2016). Exploring the mechanisms underpinning sweating: the development of a specialized ventilated capsule for use with intradermal microdialysis. *Physiol Rep* **4,** e12738.

Messick S (1987). VALIDITY. *ETS Research Report Series* **1987,** i–208.

Mitchell D & Wyndham CH (1969). Comparison of weighting formulas for calculating mean skin temperature. *Journal of Applied Physiology* **26,** 616–622.

Morris NB, Cramer MN, Hodder SG, Havenith G & Jay O (2013). A comparison between the technical absorbent and ventilated capsule methods for measuring local sweat rate. *Journal of Applied Physiology* **114,** 816–823.

Nunnally J (1978). *Psychometric theory*. McGraw Hill, New York.

Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science* **349,** aac4716–aac4716.

Patterson MJ, Stocks JM & Taylor NAS (2004). Humid heat acclimation does not elicit a preferential sweat redistribution toward the limbs. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* **286,** R512–R518.

R Development Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Relf R, Willmott A, Flint MS, Beale L & Maxwell N (2019). Reliability of a wearable sweat rate monitor and routine sweat analysis techniques under heat stress in females. *Journal of Thermal Biology* **79,** 209–217.

Shibasaki M, Aoki K, Morimoto K, Johnson JM & Takamata A (2009). Plasma hyperosmolality elevates the internal temperature threshold for active thermoregulatory vasodilation during heat stress in humans. *American Journal of*

Physiology-Regulatory, Integrative and Comparative Physiology **297,** R1706–R1712.

Shibasaki M, Wilson TE & Crandall CG (2006). Neural control and mechanisms of eccrine sweating during heat stress and exercise. *Journal of Applied Physiology* **100,** 1692–1701.

Shoukri MM, Asyali MH & Donner A (2004). Sample size requirements for the design of reliability study: review and new results. *Stat Methods Med Res* **13,** 251–271.

Shrout PE & Fleiss JL (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin* **86,** 420–428.

Taylor NA & Machado-Moreira CA (2013). Regional variations in transepidermal water loss, eccrine sweat gland density, sweat secretion rates and electrolyte composition in resting and exercising humans. *Extrem Physiol Med* **2,** 4.

Tew GA, Klonizakis M, Moss J, Ruddock AD, Saxton JM & Hodges GJ (2011). Reproducibility of cutaneous thermal hyperaemia assessed by laser Doppler flowmetry in young and older adults. *Microvascular Research* **81,** 177–182.

Vul E, Harris C, Winkielman P & Pashler H (2009). Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition. *Perspect Psychol Sci* **4,** 274–290.

TABLES

Table 1. Physical characteristics of participants (n=14).

Participant	*Age (years)*	*Height (cm)*	*Mass (kg)*	*$\dot{V}O_2$peak (mL/min/kg)*	*Body fat (%)*
1	19	170	68.8	38.4	18
2	31	177	84.5	43.7	14
3	24	178	81.9	34.0	28
4	32	181	102.3	45.6	9
5	29	184	82.7	38.5	17
6	28	170	85.2	48.9	17
7	30	168	65.1	55.0	8
8	20	179	68.8	33.0	12
9	19	177	79.1	-	-
10	18	176	66.6	48.8	16
11	22	176	70.0	61.4	-
12	23	178	98.9	37.6	21
13	19	168	70.7	40.6	20
14	21	184	94.2	38.3	10
Mean (SD)	24 (5)	176 (5)	79.9 (12.3)	43.4 (8.3)	16 (6)

Notes: Details of the screening session can be found in the Supplemental Document (https://tinyurl.com/y2wfevwz). -; data was not collected.

Table 2. Thermometry data (mean (SD)) at low-, moderate- and high-heat strain.

		Trial 1	*Trial 2*	*Trial 3*	*n1*	*n2*	*n3*
	Duration (mins)	39 (5)	38 (5)	39 (5)	14	14	14
	T_{Oeso} (°C)	37.5 (0.2)	37.5 (0.2)	37.5 (0.2)	14	14	14
	ΔT_{Oeso} (°C)	0.6 (0.1)	0.7 (0.1)	0.7 (0.1)	14	14	14
LHS	T_{Sk} (°C)	37.1 (0.4)	37.0 (0.5)	36.9 (0.3)	14	13	14
	ΔT_{Sk} (°C)	3.0 (0.5)	2.8 (0.5)	2.8 (0.5)	14	13	14
	T_B (°C)	37.4 (0.2)	37.3 (0.3)	37.4 (0.2)	14	14	14
	ΔT_B (°C)	1.1 (0.1)	1.1 (0.1)	1.1 (0.1)	14	14	14
	Duration (mins)	70 (13)	70 (9)	71 (10)	14	14	14
	T_{Oeso} (°C)	38.1 (0.3)	38.1 (0.2)	38.1 (0.2)	14	14	14
	ΔT_{Oeso} (°C)	1.2 (0.1)	1.3 (0.1)	1.3 (0.1)	14	14	14
MHS	T_{Sk} (°C)	37.7 (0.3)	37.5 (0.6)	37.6 (0.3)	14	14	14
	ΔT_{Sk} (°C)	3.6 (0.6)	3.4 (0.5)	3.5 (0.5)	14	14	14
	T_B (°C)	38.0 (0.2)	38.0 (0.2)	38.0 (0.2)	14	14	14
	ΔT_B (°C)	1.7 (0.1)	1.7 (0.1)	1.8 (0.2)	14	14	14
	Duration (mins)	107 (18)	103 (15)	104 (16)	14	14	14
	T_{Oeso} (°C)	38.6 (0.3)	38.7 (0.2)	38.7 (0.2)	14	14	14
	ΔT_{Oeso} (°C)	1.8 (0.2)	1.9 (0.1)	1.9 (0.1)	14	14	14
HHS	T_{Sk} (°C)	38.3 (0.4)	38.1 (0.5)	38.1 (0.6)	14	14	14
	ΔT_{Sk} (°C)	4.3 (0.6)	4.0 (0.6)	3.9 (0.7)	14	14	14
	T_B (°C)	38.6 (0.2)	38.5 (0.2)	38.5 (0.2)	14	14	14
	ΔT_B (°C)	2.3 (0.1)	2.3 (0.1)	2.3 (0.2)	14	14	14

Notes: LHS, low-heat strain; MHS, moderate-heat strain; HHS, high-heat strain; T_{Oeso}, oesophageal temperature; ΔT_{Oeso}, change in oesophageal temperature; T_{Sk}, mean skin temperature; ΔT_{Sk}, change in mean skin temperature; T_B, mean body temperature; ΔT_B, change in mean body temperature. n1, n2, n3, sample size in the respective trial. Mean skin temperature (T_{Sk}) was calculated as a weighted average (forehead, 7%; forearm, 14%; hand, 5%; foot, 7%; calf, 13%; quadriceps, 19%; abdomen, 35% (Mitchell & Wyndham, 1969). Mean body temperature was estimated from the weighted sum of oesophageal (0.8) and mean skin temperatures (0.2; Hardy & DuBois, 1937). Duration was calculated from the bath temperature increase to 49°C to the onset of controlled breathing (i.e., start of thermal clamp).

Table 3. Local sweat rates (LSR) (mean (SD)) of each site and region at low-, moderate- and high-heat strain, and the corresponding p-value for systematic differences.

	LSR (mg/cm^2/min)	*Trial 1*	*Trial 2*	*Trial 3*	*n1*	*n2*	*n3*	*p*
LHS	Forehead	1.04 (0.31)	0.96 (0.29)	0.90 (0.33)	11	14	14	0.462
	Chest	1.01 (0.33)	1.04 (0.32)	0.92 (0.23)	14	14	14	0.605
	Abdomen	0.82 (0.29)	0.88 (0.34)	0.81 (0.29)	14	14	12	0.897
	Bicep	0.55 (0.26)	0.67 (0.30)	0.64 (0.28)	14	14	14	0.628
	Forearm	0.77 (0.26)	0.80 (0.22)	0.81 (0.30)	14	14	14	0.960
	Hand	0.88 (0.28)	0.91 (0.33)	0.86 (0.31)	13	13	13	0.891
	Quadriceps	0.57 (0.20)	0.58 (0.18)	0.56 (0.17)	14	14	14	0.986
	Calf	0.69 (0.27)	0.69 (0.28)	0.66 (0.23)	14	14	14	0.991
	Foot	0.56 (0.30)	0.63 (0.37)	0.60 (0.29)	12	12	12	0.956
	Trunk	0.94 (0.30)	0.96 (0.23)	0.88 (0.24)	14	14	14	0.698
	Arm	0.72 (0.25)	0.77 (0.28)	0.77 (0.27)	14	14	14	0.962
	Leg	0.62 (0.23)	0.64 (0.25)	0.61 (0.21)	14	14	14	0.993
MHS	Forehead	1.44 (0.46)	1.21 (0.37)	1.26 (0.40)	11	14	14	0.362
	Chest	1.21 (0.46)	1.30 (0.34)	1.19 (0.16)	14	14	14	0.578
	Abdomen	1.03 (0.32)	1.12 (0.36)	1.06 (0.37)	14	14	12	0.808
	Bicep	0.82 (0.37)	0.88 (0.31)	0.88 (0.26)	14	14	14	0.648
	Forearm	1.05 (0.29)	0.99 (0.20)	1.04 (0.21)	13	14	14	0.863
	Hand	1.13 (0.25)	1.04 (0.31)	1.04 (0.27)	13	12	13	0.597
	Quadriceps	0.72 (0.27)	0.70 (0.19)	0.71 (0.19)	14	14	14	0.976
	Calf	0.93 (0.32)	0.84 (0.26)	0.80 (0.29)	14	14	14	0.542
	Foot	0.83 (0.38)	0.87 (0.36)	0.87 (0.33)	12	12	12	0.966
	Trunk	1.20 (0.36)	1.21 (0.25)	1.18 (0.25)	14	14	14	0.931
	Arm	0.98 (0.29)	0.95 (0.25)	0.99 (0.20)	14	14	14	0.878
	Leg	0.83 (0.29)	0.80 (0.24)	0.79 (0.23)	14	14	14	0.961
HHS	Forehead	1.45 (0.54)	1.26 (0.45)	1.37 (0.39)	11	14	14	0.584
	Chest	1.30 (0.38)	1.42 (0.36)	1.28 (0.21)	14	14	14	0.481
	Abdomen	1.09 (0.32)	1.21 (0.34)	1.12 (0.39)	14	14	13	0.446
	Bicep	0.85 (0.39)	0.94 (0.31)	0.92 (0.25)	14	14	14	0.527
	Forearm	1.06 (0.31)	1.02 (0.20)	1.06 (0.18)	14	14	14	0.919
	Hand	1.14 (0.28)	1.07 (0.29)	1.04 (0.27)	13	12	13	0.612
	Quadriceps	0.73 (0.28)	0.74 (0.21)	0.71 (0.16)	14	14	14	0.959
	Calf	1.04 (0.41)	0.87 (0.25)	0.84 (0.26)	13	14	14	0.494
	Foot	0.91 (0.40)	0.92 (0.38)	0.96 (0.36)	12	12	12	0.984
	Trunk	1.28 (0.34)	1.30 (0.26)	1.25 (0.26)	14	14	14	0.895
	Arm	1.01 (0.31)	0.99 (0.24)	1.01 (0.20)	14	14	14	0.956
	Leg	0.90 (0.31)	0.84 (0.25)	0.82 (0.22)	14	14	14	0.882

Notes: LHS, low-heat strain; MHS, moderate-heat strain; HHS, high-heat strain; Trunk: forehead, chest, abdomen; Arm: bicep, forearm, hand; Leg: quadriceps, calf, foot; n1, n2, n3, sample size in the respective trial.

Table 4. Onset threshold and thermosensitivity (mean (SD)) of each site and region, and the corresponding p-value for systematic differences.

Onset Threshold (°C)	*Trial 1*	*Trial 2*	*Trial 3*	*n1*	*n2*	*n3*	*p*
Forehead	0.4 (0.2)	0.4 (0.1)	0.4 (0.2)	10	13	14	0.848
Chest	0.4 (0.2)	0.5 (0.2)	0.4 (0.2)	13	14	14	0.800
Abdomen	0.4 (0.2)	0.4 (0.2)	0.4 (0.1)	13	14	12	0.621
Bicep	0.5 (0.2)	0.5 (0.2)	0.5 (0.2)	13	14	14	0.866
Forearm	0.5 (0.2)	0.5 (0.2)	0.5 (0.2)	13	14	14	0.908
Hand	0.5 (0.2)	0.4 (0.2)	0.4 (0.2)	12	14	13	0.828
Quadriceps	0.4 (0.2)	0.4 (0.2)	0.3 (0.2)	13	14	13	0.964
Calf	0.4 (0.2)	0.3 (0.2)	0.4 (0.2)	13	14	14	0.368
Foot	0.4 (0.2)	0.4 (0.2)	0.4 (0.2)	12	14	11	0.775
Trunk	0.4 (0.2)	0.4 (0.2)	0.4 (0.2)	13	14	14	0.931
Arm	0.5 (0.2)	0.5 (0.2)	0.4 (0.2)	13	14	14	0.919
Leg	0.4 (0.2)	0.4 (0.2)	0.4 (0.2)	13	14	14	0.870
Thermosensitivity (mg/cm^2/min/°C)	*Trial 1*	*Trial 2*	*Trial 3*	*n1*	*n2*	*n3*	*p*
Forehead	1.75 (0.85)	1.44 (0.63)	1.56 (0.77)	10	13	13	0.607
Chest	1.89 (0.77)	1.88 (0.81)	1.93 (1.13)	13	14	14	0.949
Abdomen	1.76 (0.77)	1.65 (0.75)	1.52 (0.87)	13	14	12	0.462
Bicep	1.32 (0.61)	1.28 (0.50)	1.09 (0.54)	13	14	14	0.530
Forearm	1.67 (0.84)	1.51 (0.78)	1.56 (0.84)	13	14	14	0.828
Hand	1.90 (0.51)	1.41 (0.71)	1.69 (0.81)	12	14	13	0.135
Quadriceps	1.03 (0.49)	1.07 (0.85)	1.04 (0.60)	13	14	13	0.980
Calf	1.31 (0.57)	1.06 (0.52)	1.19 (0.53)	13	14	14	0.555
Foot	1.06 (0.46)	0.87 (0.48)	0.91 (0.48)	12	14	11	0.502
Trunk	1.82 (0.72)	1.66 (0.57)	1.68 (0.90)	13	14	14	0.683
Arm	1.60 (0.63)	1.40 (0.60)	1.44 (0.69)	13	14	14	0.662
Leg	1.15 (0.44)	1.00 (0.55)	1.06 (0.52)	13	14	14	0.567

Notes: Trunk: forehead, chest, abdomen; Arm: bicep, forearm, hand; Leg: quadriceps, calf, foot; n1, n2, n3, sample size in the respective trial.

Table 5. A table based on the reliabilty statistics determined from the current study to aid researchers in making a decision of which site or region is most approporiate to measure local sweat rate based on level of heating and sites available.

	LHS (Δ0.6°C)		**MHS (Δ1.2°C)**		**HHS (Δ1.8°C)**	
	CV < 25%	ICC ≥ 0.70	CV < 25%	ICC ≥ 0.70	CV < 25%	ICC ≥ 0.70
Forehead			✓			
Chest	✓				✓	
Abdomen	✓	✓	✓	✓	✓	✓
Forearm	✓	✓	✓		✓	
Hand	✓	✓	✓	✓	✓	
Quadriceps	✓	✓	✓	✓	✓	✓
Calf	✓	✓			✓	✓
Foot		✓		✓	✓	✓
Trunk	✓	✓	✓		✓	✓
Arm	✓	✓	✓		✓	
Leg	✓	✓	✓	✓	✓	✓

***Notes*:** LHS, low-heat strain (pink); MHS, moderate-heat strain (teal); HHS, high-heat strain (purple); level of heat strain refers to changes in oesophageal temperature (for exact changes, see Table 2); Trunk: forehead, chest, abdomen; Arm: bicep, forearm, hand; Leg: quadriceps, calf, foot; CV, coefficient of variation; ICC, intraclass correlation coefficient. Checkmarks indicate which sites meet the thresholds for absolute and/or relative reliability. A CV criterion of 25% (moderate absolute reliability) is used since no sites demonstrated good absolute reliability (CV ≤10%).

FIGURE LEGENDS

Figure 1. A) Visual representation of measurement locations (created using BioRender.com) for local sweat rates with the corresponding landmarking instructions used for the placement of the capsules and colour indicating location on graph in Panel B. Created using BioRender.com **B)** Representative profile of sweating demonstrating the periods used in data analysis. LSR, local sweat rate; ΔT_{Oeso}, change in oesophageal temperature (black line); NHS, no-heat strain; LHS, low-heat strain; MHS, moderate-heat strain; HHS, high-heat strain.

Figure 2. Local sweat rate (LSR) averaged over the three experimental trials (Panel A), intraclass correlation coefficient (ICC; Panel B), coefficient of variation (CV; Panel C) and minimum detectable change (MDC; Panel D) at each site and region (trunk: forehead, abdomen, chest; arm: bicep, forearm, hand; leg: quadriceps, calf, foot) during whole body passive heating at low-heat strain. LSR is presented as mean (SD). Reliability statistics are presented as mean and 95% CI. Dotted lines indicate ICC ≥0.70 (Panel B) and CV <25% (Panel C).

Figure 3. Local sweat rate (LSR) averaged over the three experimental trials (Panel A), intraclass correlation coefficient (ICC; Panel B), coefficient of variation (CV; Panel C) and minimum detectable change (MDC; Panel D) at each site and region (trunk: forehead, abdomen, chest; arm: bicep, forearm, hand; leg: quadriceps, calf, foot) during whole body passive heating at moderate-heat strain. LSR is presented as mean (SD). Reliability statistics are presented as mean and 95% CI. Dotted lines indicate ICC ≥0.70 (Panel B) and CV <25% (Panel C).

Figure 4. Local sweat rate (LSR) averaged over the three experimental trials (Panel A), intraclass correlation coefficient (ICC; Panel B), coefficient of variation (CV; Panel C) and minimum detectable change (MDC; Panel D) at each site and region (trunk: forehead, abdomen, chest; arm: bicep, forearm, hand; leg: quadriceps, calf, foot) during whole body passive heating at high-heat strain. LSR is presented as mean (SD). Reliability statistics are presented as mean and 95% CI. Dotted lines indicate ICC ≥0.70 (Panel B) and CV <25% (Panel C).

Figure 5. Onset threshold (Panel A) and Thermosensitivity (Panel B) averaged over the three experimental trials, intraclass correlation coefficient (ICC), coefficient of variation (CV) and minimum detectable change (MDC) at each site and region (trunk: forehead, abdomen, chest; arm: bicep, forearm, hand; leg: quadriceps, calf, foot) during whole body passive heating. Onset threshold and thermosensitivity are presented as mean (SD). Reliability statistics are presented as mean and 95% CI. Dotted lines indicate ICC ≥0.70 and CV <25%.

Figure 6. Effect size of the incremental whole body passive heating model on sweat rate at each site and region during each trial when considered from: A) no-heat strain (NHS) to low-heat strain (LHS); B) NHS to moderate-heat strain (MHS); and C) NHS to high-heat strain (HHS).

Figure 7. Example Scenario: A researcher is interested in measuring a change in local sweat rate (LSR) at the forearm in response to an acclimation protocol. The anticipated change (~25%) relates to ~0.26 mg/cm^2/min, and a standardised effect size of 1.083 (Cohen's *d*). As a conservative estimate, the researcher considers the minimum detectable change (MDC) calculated using the absolute reliability statistics and expressed relative to the 70% confidence level. The standardised effect size of this MDC (*d*: 1.042) represents minimal difference between their anticipated effect and a level of confidence they consider appropriate for their circumstance (~0.01 mg/cm^2/min) to the anticipated effect, and an acceptable confidence level for their purposes. The researchers identify the value on the effect size panel (**1**) and then the respective power curve (**2**). By tracking across the point that the power curve hits 0.8 (representing 80% power, or a 0.2 beta), they identify that 10 participants are required for this within-subject design using a simple t-test and 0.05 critical alpha (**3**).

Figure 1.

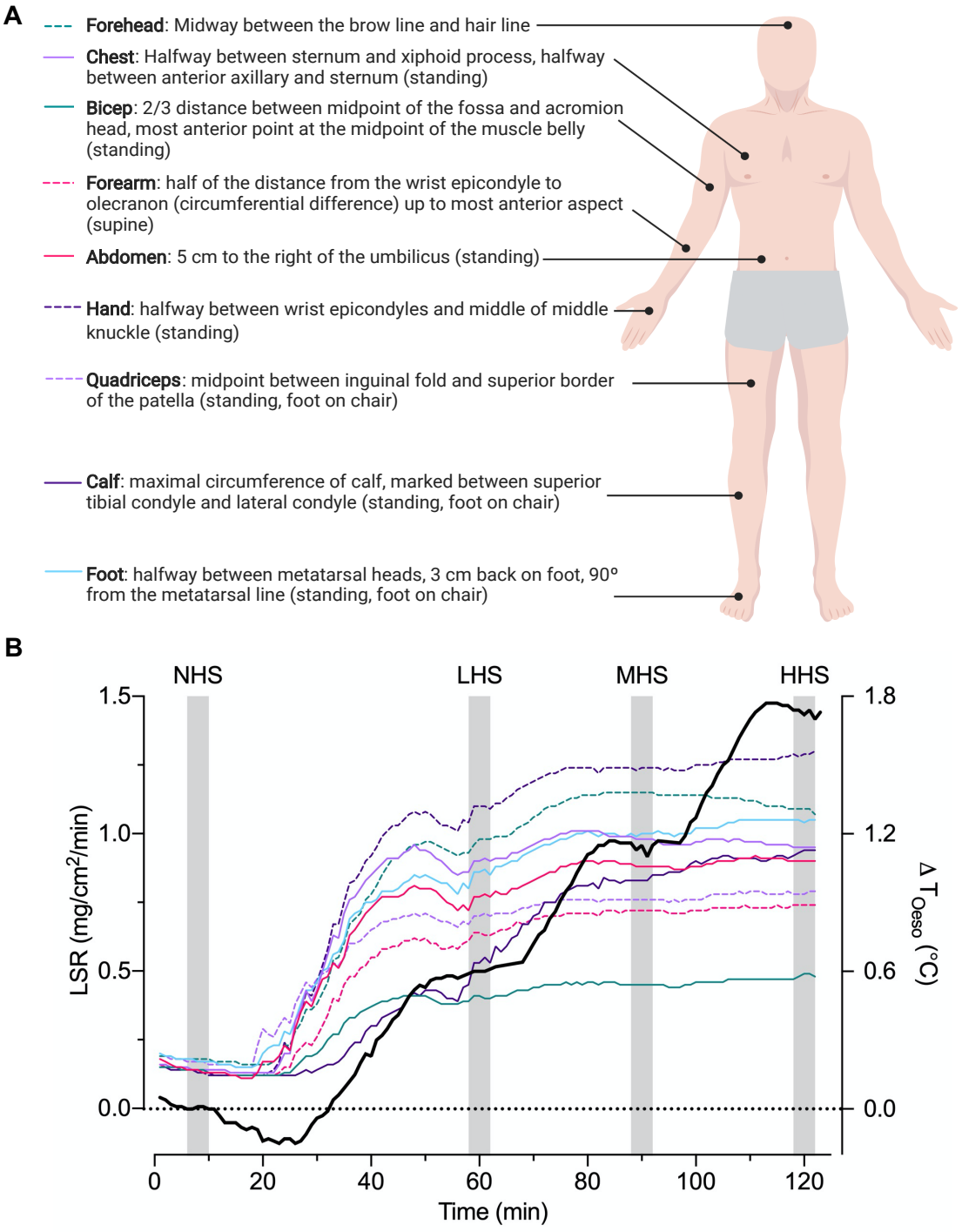

A
Forehead: Midway between the brow line and hair line
Chest: Halfway between sternum and xiphoid process, halfway between anterior axillary and sternum (standing)
Bicep: 2/3 distance between midpoint of the fossa and acromion head, most anterior point at the midpoint of the muscle belly (standing)
Forearm: half of the distance from the wrist epicondyle to olecranon (circumferential difference) up to most anterior aspect (supine)
Abdomen: 5 cm to the right of the umbilicus (standing)
Hand: halfway between wrist epicondyles and middle of middle knuckle (standing)
Quadriceps: midpoint between inguinal fold and superior border of the patella (standing, foot on chair)
Calf: maximal circumference of calf, marked between superior tibial condyle and lateral condyle (standing, foot on chair)
Foot: halfway between metatarsal heads, 3 cm back on foot, 90° from the metatarsal line (standing, foot on chair)
B
NHS
LHS
MHS
HHS
LSR (mg/cm2/min)
ΔTOeso (°C)
Time (min)

Figure 2.

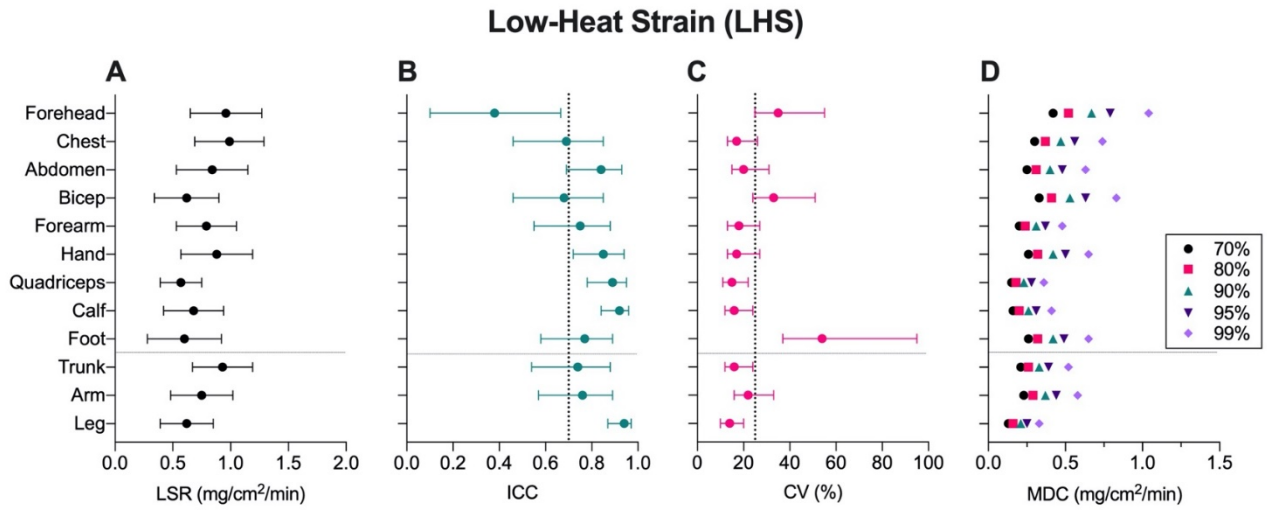

Low-Heat Strain (LHS)
A
B
C
D
Forehead
Chest
Abdomen
Bicep
Forearm
Hand
Quadriceps
Calf
Foot
Trunk
Arm
Leg
0.0 0.5 1.0 1.5 2.0
LSR (mg/cm2/min)
0.0 0.2 0.4 0.6 0.8 1.0
ICC
0 20 40 60 80 100
CV (%)
0.0 0.5 1.0 1.5
MDC (mg/cm2/min)
70%
80%
90%
95%
99%

Figure 3.

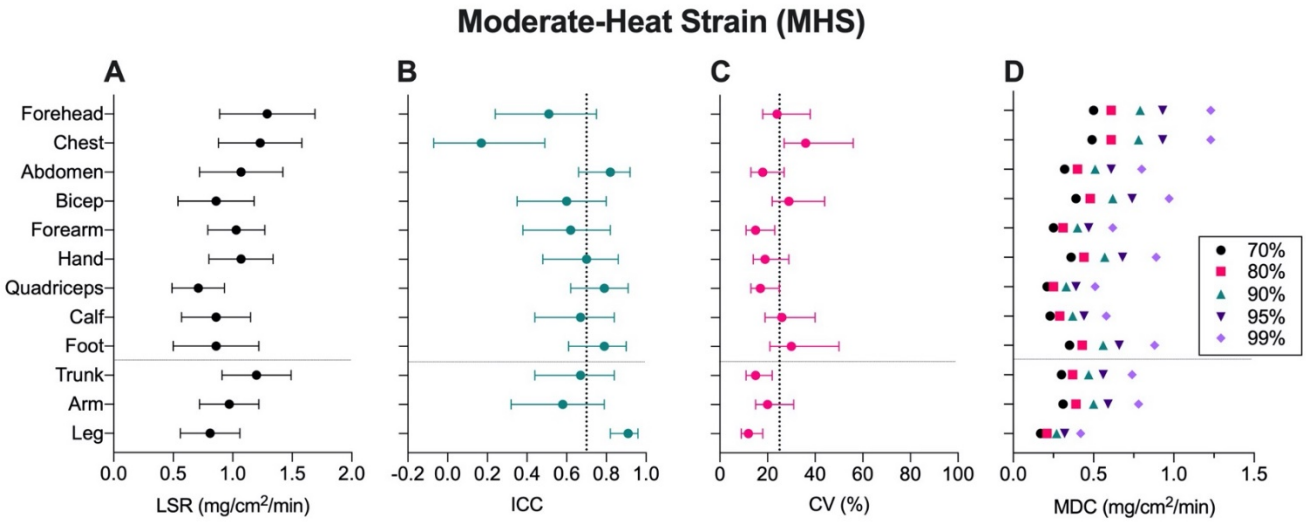

Moderate-Heat Strain (MHS)
A
B
C
D
Forehead
Chest
Abdomen
Bicep
Forearm
Hand
Quadriceps
Calf
Foot
Trunk
Arm
Leg
0.0 0.5 1.0 1.5 2.0
LSR (mg/cm2/min)
-0.2 0.0 0.2 0.4 0.6 0.8 1.0
ICC
0 20 40 60 80 100
CV (%)
0.0 0.5 1.0 1.5
MDC (mg/cm2/min)
70%
80%
90%
95%
99%

Figure 4.

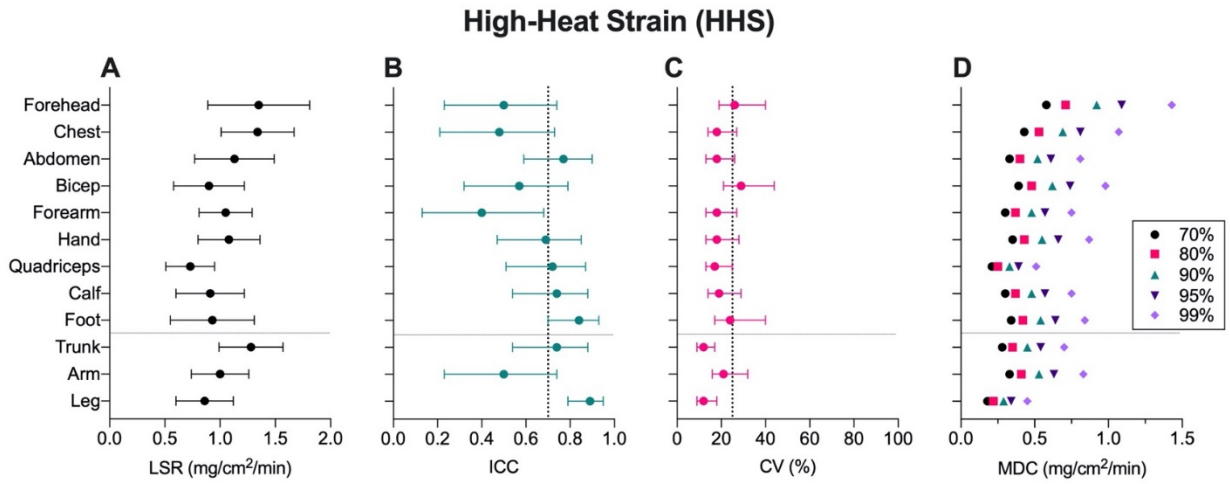

High-Heat Strain (HHS)
A
B
C
D
Forehead
Chest
Abdomen
Bicep
Forearm
Hand
Quadriceps
Calf
Foot
Trunk
Arm
Leg
0.0 0.5 1.0 1.5 2.0
LSR (mg/cm2/min)
0.0 0.2 0.4 0.6 0.8 1.0
ICC
0 20 40 60 80 100
CV (%)
0.0 0.5 1.0 1.5
MDC (mg/cm2/min)
70%
80%
90%
95%
99%

Figure 5.

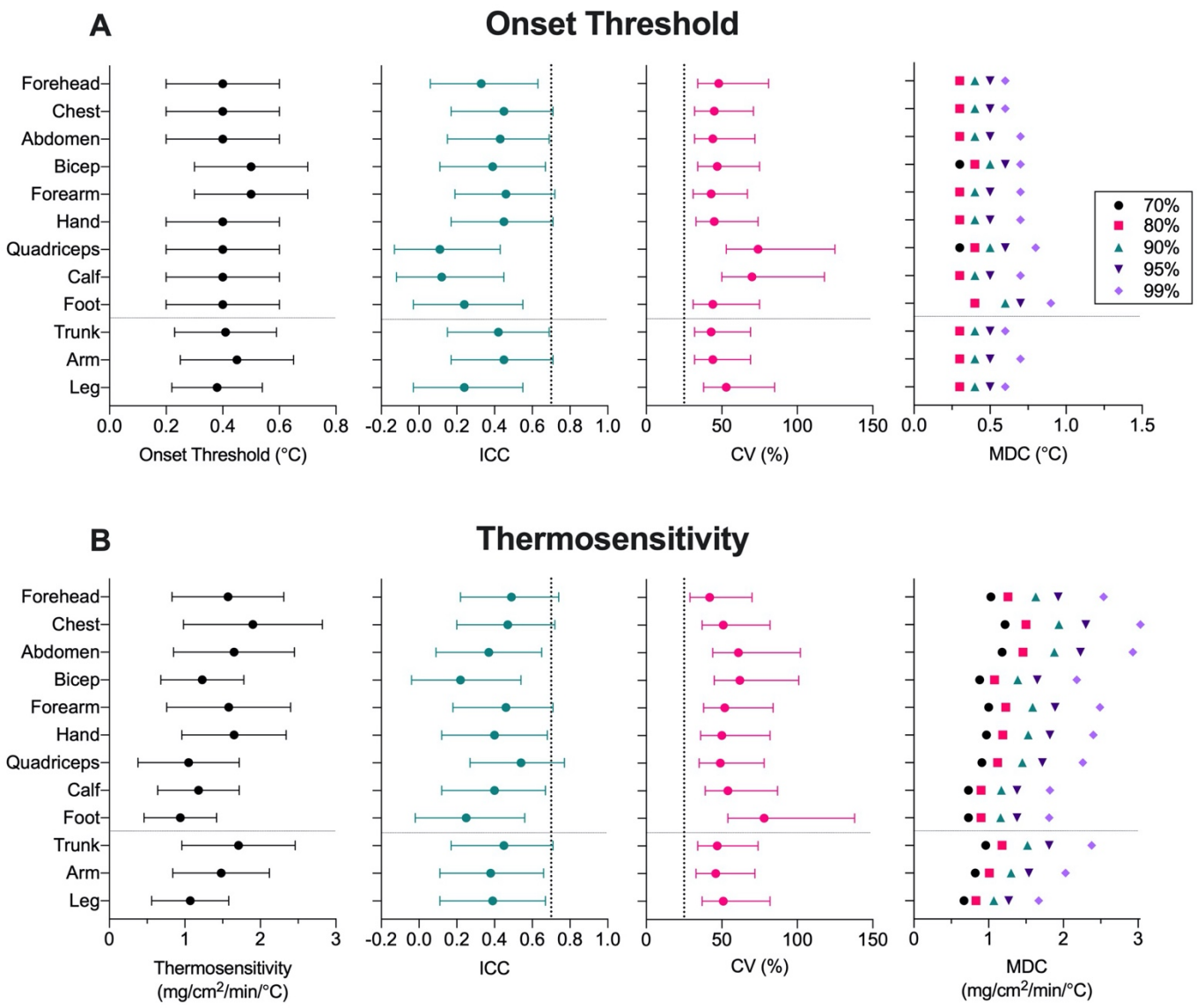

A
Onset Threshold
Forehead
Chest
Abdomen
Bicep
Forearm
Hand
Quadriceps
Calf
Foot
Trunk
Arm
Leg
0.0 0.2 0.4 0.6 0.8
Onset Threshold (°C)
-0.2 0.0 0.2 0.4 0.6 0.8 1.0
ICC
0 50 100 150
CV (%)
0.0 0.5 1.0 1.5
MDC (°C)
70%
80%
90%
95%
99%
B
Thermosensitivity
0 1 2 3
Thermosensitivity
(mg/cm2/min/°C)
-0.2 0.0 0.2 0.4 0.6 0.8 1.0
ICC
0 50 100 150
CV (%)
0 1 2 3
MDC
(mg/cm2/min/°C)

Figure 6.

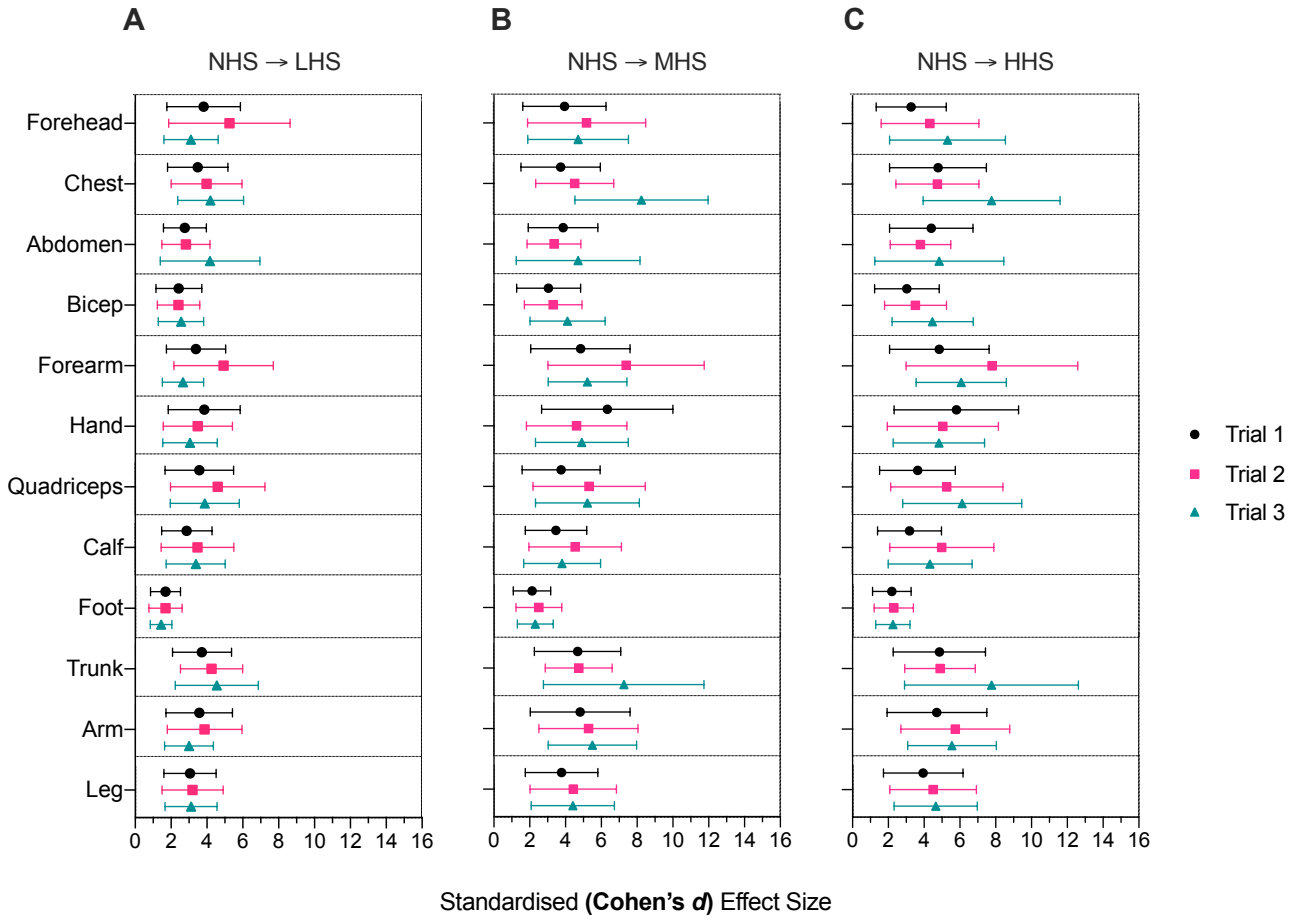
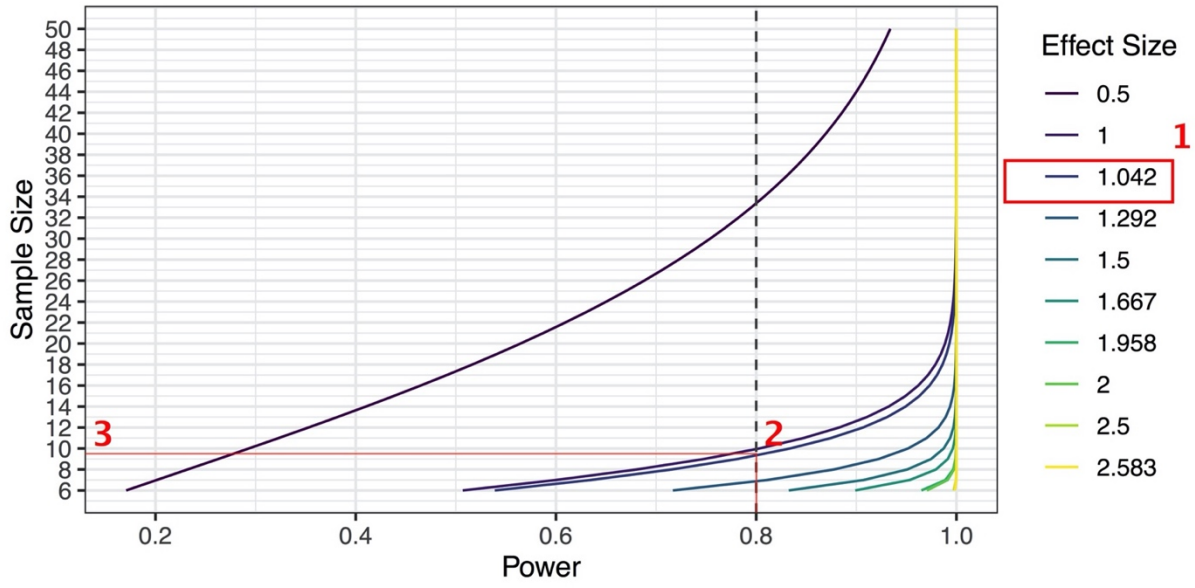

A
NHS → LHS
B
NHS → MHS
C
NHS → HHS
Forehead
Chest
Abdomen
Bicep
Forearm
Hand
Quadriceps
Calf
Foot
Trunk
Arm
Leg
0 2 4 6 8 10 12 14 16
Trial 1
Trial 2
Trial 3
Standardised (Cohen's d) Effect Size

Figure 7.


Sample Size
50
48
46
44
42
40
38
36
34
32
30
28
26
24
22
20
18
16
14
12
10
8
6
3
2
0.2
0.4
0.6
0.8
1.0
Power
Effect Size
0.5
1
1.042
1.292
1.5
1.667
1.958
2
2.5
2.583
1

PART THREE: EXTENDED DISCUSSION

3. Extended Discussion

3.1 Elaboration on the Decision Table

The decision table (Table 5) was created to facilitate the selection of the most appropriate site for measuring sweating will be a valuable tool for researchers. Absolute and relative measures of reliability can have independent and combined applications to research, hence the inclusion of both criteria separately in the decision table. Absolute reliability, as indicated by a coefficient of variation (CV) less than 25%, is the amount of error (biological, measurement, etc.) associated with a given measurement. Since the coefficient of variation is the typical error relative to the mean, it enables the coefficient of variation to be compared across measurement techniques or populations. Studies aiming to determine the effect of an intervention between groups (young vs older adults, male vs female), would be better suited to a site that is reliable based on CV. Since coefficient of variation is represented as error relative to the mean, it would then be appropriate to be applied to both populations being studied. Additionally, given that the minimum detectable changes are based on the typical error (coefficient of variation), it will enable researchers to determine whether a change is beyond the noise.

The intraclass correlation coefficient (ICC) is a measure of relative reliability, or the consistency of a measurement's rank within a sample. Studies assessing individual differences would benefit from selecting a site based on relative reliability. For example, if a researcher wanted to determine why an individual factor such as fitness impacted local sweat rate, relative reliability will be more important to help differentiate among a group of individuals. However, if the sample of individuals is homogenous and all the participants have very similar fitness levels, it would be difficult to distinguish among these

individuals since even the smallest change in sweating could affect the rank of multiple subjects. When samples are more heterogenous (i.e., a greater range of fitness levels) it will be easier to discern among individuals as even a large change is unlikely disrupt the rank order. Relative reliability is also helpful for within group analysis, to determine why individuals may be changing position within a group. For this reason, studies that have a within and between group design (i.e. determining the effect of heat acclimation on sweating in young vs older adults) may consider using a site that is reliable by both absolute and relative standards to assist with the within (relative) and between (absolute) group analysis.

At this time, what a meaningful change in sweating is in terms of heat loss ability is has not been explicitly established. Once it is determined how much of a change in local sweat rate contributes to a change in core temperature, what would be a meaningful difference in sweating and what would be considered 'good' reliability can be determined.

3.2 Regional Variations in Sweating

Although this study was not designed to compare sweat rates between regions (i.e., low power for this purpose), these descriptive findings warrant discussion in light of previous research. The trunk region consistently exhibited the highest sweat rate throughout each level of heating. Sweat rate at the arm was ~ 80% the rate of the trunk, while the leg had the lowest sweat rates, equalling ~84% that of the arm. This regional distribution of sweat rates is similar to previous reported of higher sweating rates at the torso and forehead, and lower sweat rates at the periphery (Cotter, Patterson & Taylor, 1995; Smith & Havenith, 2011, 2012). These sites have the largest surface area, and

therefore the greatest potential for evaporation as highlighted by Taylor and Machado-Moreira (2013), based on high sweating rates at the back, thigh, head, leg, chest and abdomen. Consistent with these findings, the abdomen, chest and forehead exhibited the highest sweat rates throughout the passive heating protocol (Figures 2-4). However, to the contrary, the quadriceps (thigh) exhibited the lowest sweat rate; a discrepancy which could be due to different models used to stimulate sweating (i.e., passive vs. exercise; Taylor & Machado-Moreira, 2013). The advantage of whole-body passive heating is the ability to clamp both core and skin temperatures, whereas exercise models are subject to regional variations in skin temperature. In particular, high lower-body skin temperatures due to heat production from active muscle tissue can result in regional variations in sweating, particularly given the large muscle mass in the thigh (Kenny, 2014; Todd, Gordon, Groeller & Taylor, 2014). Additionally, exercise stimulates sweating both thermally and non-thermally (Van Beaumont & Bullard, 1963), a factor that may contribute to the discrepancies with the current findings.

Interestingly, there were minimal increases in sweat rate (≤0.05 mg/cm^2/min) at the peripheral regions from moderate- to high-heat strain (with the exception of the foot), while the trunk region (forehead, abdomen and chest) continued to increase. Sweat rate is dependent first on the recruitment and activation of more sweat glands, then the output from these glands (Kondo et al. 2001). There are regional variations in the density (Taylor & Machado-Moreira, 2013), size (Sato & Sato, 1983) and innervation of sweat glands (Kondo et al. 1998), which likely explains differences in sweating profile between sites. This is particularly likely given that the recruitment of activated sweat glands is saturated within approximately 8 minutes of exercise or passive heat exposure, and further

increases in sweat rate are due to changes in sweat gland output (Kondo et al. 2001). Collectively, this implies that the sites which exhibited a plateau in sweat rate (bicep, forearm, hand, quadriceps, calf), reached maximal recruitment of sweat glands and maximal sweat gland output at lower levels of heat stress compared to the chest, abdomen, forehead and foot, which presented continued increases in sweat gland output throughout heating. This highlights the contribution that measuring sweat gland activation could provide to interpreting results in future studies similar to this. Additionally, the flow rate in the current study (0.75 L/min) was constant throughout the heating protocol, so this occurrence may also result from methodological constraints. Although pilot testing was completed to select a flow rate that would ensure full sweat evaporation at the highest level of heat strain, it may not have been high enough to facilitate complete evaporation at regions with high sweat rates among individuals who produce high volumes of sweat, resulting in an artificial increase in sweating.

3.3 Onset Threshold

This study was the first to determine the onset threshold of sweating at nine sites on the body, extending our previous understanding which was restricted to the forearm (Brengelmann, Savage, & Avery, 1994; Kenefick et al. 2012). Previously, smaller variations in onset threshold compared to sweat rate led researchers to conclude that onset threshold is a more reliable measure (Kenefick et al. 2012). This observation, in combination with the large effect of interventions on onset threshold (hyperosmolality, sex, age: Barrera-Ramirez, McGinn, Carter, Franco-Lopez, & Kenny, 2014; Smith, Alexander & Kenney, 2013), prompted the suggestion that onset threshold may be better

indicator of thermoeffector function than local sweat rate (Kenefick et al. 2012). The current study extends the knowledge of reliability of onset thresholds across the body, given that the forearm may not be accessible for some study designs. However, all sites demonstrated poor relative (ICC <0.70) and absolute (CV >25%) reliability. Caution must be used with the use and interpretation of onset threshold reliability, as the analysis and determination of onset threshold is dependent on researcher interpretation. Although in the current study the researcher who performed the analysis was blinded therefore reducing (but not removing) bias, optimal measures of reliability come from data that does not require researcher interpretation.

The recruitment pattern of sweating has been subject to continued debate, initially thought to be a caudal to rostral pattern (Hertzman et al. 1952; Park & Tamura, 1992; Rawson & Randall, 1961), more recent work has shown the pattern appears to be random (Frei et al. 2019). The current study design provides insight into the recruitment pattern across nine sites. When assessed qualitatively, the quadriceps was the first site to start sweating the most, followed by the calf site, which was the second to be activated (Table 7). The abdomen, calf, foot and quadricep were all activated third the same number of times. The abdomen was the site most often activated fourth, and fifth in tandem with the hand. The chest was activated the most frequently as sixth, the bicep was seventh, the forearm was most frequently eighth and the foot was most often last. When considering average position, or the place in the recruitment pattern at which a site was most frequently activated, all sites were either third, fifth or sixth, lending support that recruitment is random as there is no discernable pattern (Table 7). Of note, the forehead, quadricep and calf were the three most common sites to be activated first and second,

refuting the caudal-rostral pattern which suggests the onset of sweating begins distally and proceeds upwards (Hertzman et al. 1952). Further, the chest, hand, abdomen and foot all exhibited the same average rank, further disputing that proposition. Although the activation pattern was not a primary outcome measure of the current study, it furthers our understanding of the recruitment of sweating.

Table 7. The number of times a site was activated first, second, third, fourth, fifth, sixth, seventh, eighth and ninth in the activation sequence and the average activation position across participants at each site.

	Abdomen	Bicep	Calf	Chest	Foot	Forearm	Forehead	Hand	Quadriceps
Times site was activated first	2	0	6	3	3	1	8	4	10
...second	3	2	8	2	2	2	7	1	6
...third	4	0	4	3	4	2	1	2	4
...fourth	5	2	2	4	4	0	3	2	2
...fifth	4	2	0	2	2	2	2	4	1
...sixth	1	6	2	8	1	4	2	4	0
...seventh	3	8	0	3	1	4	1	5	1
...eighth	2	5	3	0	0	6	2	3	1
...ninth	2	1	1	1	9	5	0	1	1
Average Activation Position	5	6	3	5	5	6	3	5	3

3.4 Thermosensitivity

Previous studies have assessed regional heterogeneity in thermosensitivity during exercise, reporting higher values at the forehead, chest, abdomen and hand, and lower values at the forearm, thigh and foot (Aoki et al. 1995; Kondo et al. 1998; Machado-Moreira, Caldwell, Mekjavic, & Taylor, 2008; Machado-Moreira, Smith, van den Heuvel, Mekjavic, & Taylor, 2008; Machado-Moreira, Wilmink, Meijer, Mekjavic, & Taylor, 2008; Taylor, Caldwell, & Mekjavic, 2006). Although this pattern does not come from simultaneous measurements but rather from a series of studies, the current study found a similar trend in regional thermosensitivity (with the exception of the forearm), reporting

higher values at the forehead, chest, abdomen, forearm and hand, and lower values at the bicep, quadriceps (thigh), calf and foot (Figure 5B). Despite similar patterns in sudomotor thermosensitivity, actual values differed between studies, likely due to differences in the heating model (passive heating vs. exercise) as discussed previously. Further, as this study is the first to assess the reliability of thermosensitivity, the unreliability of the response itself may be the cause of discrepancies between the studies.

3.5 Future Directions

3.5.1 *Inter-individual Factors*

The limitations presented of the current thesis offer many future directions of research. Investigations into sex-related differences in sweating have demonstrated that although males tend have higher sweat rates than females (Smith & Havenith, 2012), menstrual cycle phase does not appear to modulate heat loss (Notley, Dervis, Poirier, & Kenny, 2019). However, it has been well documented that menstrual cycle alters the onset threshold of sweating in women, although only assessed at one skin site (arm) or via whole body sweat rate (Bittel & Henane, 1975; Stephenson & Kolka, 1985, 1999). Other studies determined that the differences in onset of sweating between phases was attributable to differences in resting core temperature caused by changing hormone concentrations (Frascarolo, Schutz, & Jequier, 1992; Haslag & Hertzman, 1965). Women have also been shown to have a lower thermosensitivities at the forearm than their male counterparts (Barrera-Ramirez et al. 2014). As discussed previously, the aim of reliability studies is to quantify the error (biological, measurement, or other) of a given measurement. Biological error can stem from many sources, including menstrual cycle. Given the physiological variability that accompanies the fluctuations in hormones

throughout the menstrual cycle, it is likely that the sweating response is more variable both within and between female individuals. It is possible that meaningful changes in sweating are phase-dependent, in addition to being site and heat load dependent. Establishing this would contribute and facilitate to future research on females, through being able to compare between phases by having the error or variability accounted for in the measurement. Fluctuating hormones can make it difficult for researchers to create similar trial conditions and is still not fully understood as to how it may affect different aspects of thermoregulation. Determining the reliability of local sweating across phases of the menstrual cycle would contribute greatly to the field of environmental physiology.

In the current study, it was anticipated that regional variations in sweating would extend to regional variations in reliability. Older adults have reduced sweating compared to their younger counterparts, although fitness can modify this response, and this is region dependent, meaning there are regional variations in the age-related decrement in sweating (Inoue, Nakao, Araki, & Murakami, 1991; Inoue & Shibasaki, 1996; Smith, Alexander & Kenney, 2013). Thus, it is likely that for the same reasons as the current study, reliability of local sweat rate will be different among older adults. Local sweat rate is often used to assess thermoeffector responses and is used to explore age-related sweating impairments, thus it is important to extend our knowledge of reliability to older adults. Further, as this study is the first to establish sweat rate at three levels of heat stress, onset and thermosensitivity at nine regions of the body in young men, it remains unknown in older adults. Examining local sweating in older adults using a similar model to the current study would help to determine whether the age-related decrement in

sweating is not only dependent on region, but also heating level and therefore level of heat strain.

Similar to older adults, individuals with chronic diseases such as diabetes or hypertension warrant their own investigation into regional variations in sweating and reliability as these populations have diminished thermoregulatory abilities (Kenny, Yardley, Brown, Sigal, & Jay, 2010). Specifically, impairments in local sweating has been shown in individuals with diabetes, typically exhibiting lower body reductions and upper body increases in sweat rate (Carter, McGinn, Barrera-Ramirez, Sigal & Kenny, 2014; Petrofsky, Lee, Patterson, Cole, & Stewart, 2005). It is believed that this regional pattern of sweating stems from diabetes-related neuropathies that begin distally (Fealey, Low, & Thomas, 1989; Hoeldtke, Bryner, & VanDyke, 2011; Petrofsky, Berk, & Al-Nakhli, 2012). Previous research from our laboratory determined that although the reductions in sweat rate among individuals with type 1 diabetes were similar across regions, the impairments stemmed from different mechanistic restrictions such that the chest was reduced due to fewer active sweat glands while the forearm was due to reduced output (Carter et al. 2014). Further, the same study determined that impairments in heat loss amongst individuals with type 1 diabetes is heat load dependent. This emphasises the importance of assessing regional dependent impairments in a model such as the one in the current study, to determine if regional impairments in sweating are heat load dependent, but also if it affects the reliability of these responses as well. Given that changes in sudomotor responses are indicative of sympathetic dysfunction (Hoeldtke, Bryner, Horvath, Phares, Broy, & Hobbs, 2001; Luo, Chao, Hsieh, Lue, & Hsieh, 2012), knowing the reliability, and therefore the meaningful change of sweating responses across various regions of the

body (especially lower body), is critical to the monitoring and treatment of these individuals.

3.5.2 Intra-individual Factors

Dehydration is common during heat strain, particularly during prolonged exercise in the heat, as fluid consumption is often inadequate to replace the high volume of fluid lost via sweating (Sawka, Toner, Francesconi, & Pandolf, 1983; McKinley, Martelli, Pennington, Trevaks, & McAllen, 2018). As the body loses fluid via sweat (which is hypotonic to blood plasma), plasma osmolality increases creating hypovolemic hyperosmotic conditions, also known as hypohydration (McKinley & Johnson, 2004). Hypohydration impairs sweat production (thereby compromising evaporative heat loss) firstly through the onset of sweating. It was determined that the change in core (esophageal) temperature required to initiate sweating corresponds to changes in plasma osmolality, with higher levels of hyperosmolality having higher onset temperatures (Takamata, Nagashima, Nose, & Morimoto, 1997). Similar findings were shown by Shibasaki, Aoki, Morimoto, Johnson, and Takamata (2009), in hypohydrated participants during passive heating where the onset threshold for sweating was 36.63°C and 37.73°C for isosmotic and hyperosmotic conditions respectively.

Hypohydration also reduces sweat rate during exercise; previous studies have shown that regardless of larger increases in core temperature when hypohydrated compared to euhydrated, the sweating response is reduced when hypohydrated during heat strain (Greenleaf & Castle, 1971; Sawka et al. 1983). However, the reduction in sweating is not uniform across the body with one study demonstrating reductions in sweat

rate at the forehead but not the scapula (Machado-Moreira et al. 2009). While whole-body sweating has been the subject of extensive research, there is considerably less known about the interactive effects of hydration status on regional differences in sweat rate, onset threshold and thermosensitivity, and the reliability of these responses.

Non-specific to the field of thermal physiology, determining what a meaningful change is among women, older adults, individuals with chronic conditions and between physiological states has important clinical implications. Whether it be for choices related to a diagnosis or treatment of these populations, it is imperative that those decisions are made off of meaningful changes.

PART FOUR: GENERAL CONCLUSIONS OF THE THESIS

4. General Conclusions of the Thesis

The current thesis was the first to determine the reliability of local sweat rate at nine sites and three regions of the body during low-, moderate- and high-heat strain. Use of the ventilated capsule technique allowed for the determination of the reliability of the onset threshold and thermosensitivity at each of these sites, expanding the novelty of this study. The information can help researchers in calculating sample size, selecting the most appropriate site of measurement and data interpretation.

The primary implication is that the reliability of local sweat rate depends on site or region of measurement and level of heat strain as it appears that reliability of sweating is inversely related to level of heat strain. Reliability was assessed based on sites meeting criteria of acceptable relative reliability (ICC ≥0.70), moderate absolute reliability (CV <25%), or both. At low-heat strain, most sites and all regions demonstrated acceptable relative reliability and moderate absolute reliability. At moderate-heat strain, the abdomen, hand, quadriceps, foot and leg had acceptable relative reliability while the forehead, abdomen, forearm, hand, quadriceps and all three regions had moderate absolute reliability. At high-heat strain, relative reliability was acceptable at the abdomen, quadriceps, calf, foot and leg. Absolute reliability was moderate at the chest, abdomen, forearm, hand, quadriceps, calf, foot and all regions. Neither onset threshold nor thermosensitvity demonstrated acceptable relative or moderate absolute reliability at any site. However, it is important to consider that both variables are subject to researcher interpretation, which may in turn, affect the reliability of these measures.

The findings of the current study reinforce that researchers should select a site or region of measurement based on their intended heating level and what sites are available

to measure sweating. My thesis provides a framework to help future researchers make these decisions (Table 5) and think critically about their research design based on the intended purpose and expected findings. By providing both the absolute and relative reliability of each site, researchers will be able to select a measurement site or region based on the reliability most appropriate for their study design. Further, by presenting the minimum detectable changes for each site and variable, researchers can use this to interpret their own data and determine their confidence that a given change is due to their intervention, rather than error. Finally, the current thesis provides a thought-provoking foundation for the continued research into the regional variations of sweat rate, onset threshold, thermosensitivity and the reliability of these responses among different populations and physiological conditions.

PART FIVE: REFERENCES

5. References

Aoki, K., Shiojiri, T., Shibasaki, M., Takano, S., Kondo, N., & Iwata, A. (1995). The effect of diurnal variation on the regional differences in sweating and skin blood flow during exercise. *European journal of applied physiology and occupational physiology*, *71*(2-3), 276-280.

Atkinson, G., & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports medicine*, *26*(4), 217-238.

Baker, L. B., Barnes, K. A., Anderson, M. L., Passe, D. H., & Stofan, J. R. (2016). Normative data for regional sweat sodium concentration and whole-body sweating rate in athletes. *Journal of sports sciences*, *34*(4), 358-368.

Barrera-Ramirez, J., McGinn, R., Carter, M. R., Franco-Lopez, H., & Kenny, G. P. (2014). Osmoreceptors do not exhibit a sex-dependent modulation of forearm skin blood flow and sweating. *Physiological reports*, *2*(2), e00226.

Beaton, D. E. (2000). Understanding the relevance of measured change through studies of responsiveness. *Spine*, *25*(24), 3192-3199.

Benzinger TH. (1969). Heat regulation: homeostasis of central temperature in man. *Physiological Reviews* **49,** 671-759.

Bittel, J., & Henane, R. (1975). Comparison of thermal exchanges in men and women under neutral and hot conditions. *The Journal of physiology*, *250*(3), 475-489.

Brengelmann, G. L., McKeag, M., & Rowell, L. B. (1975). Use of dew-point detection for quantitative measurement of sweating rate. *Journal of applied physiology*, *39*(3), 498-500.

Brengelmann, G. L., Savage, M. V., & Avery, D. H. (1994). Reproducibility of core temperature threshold for sweating onset in humans. *Journal of Applied Physiology*, *77*(4), 1671-1677.

Buley, H. M. (1938). Active sweat glands: a method for their study. *Archives of Dermatology and Syphilology*, *38*(3), 340-348.

Cain B & McLellan TM. (1998). A model of evaporation from the skin while wearing protective clothing. *Int J Biometeorol* **41,** 183-193.

Carter, M. R., McGinn, R., Barrera-Ramirez, J., Sigal, R. J., & Kenny, G. P. (2014). Impairments in local heat loss in type 1 diabetes during exercise in the heat. *Medicine & Science in Sports & Exercise*, *46*(12), 2224-2233.

Cheuvront, S. N., Bearden, S. E., Kenefick, R. W., Ely, B. R., DeGroot, D. W., Sawka, M. N., & Montain, S. J. (2009). A simple and valid method to determine thermoregulatory sweating threshold and sensitivity. *Journal of Applied Physiology*, *107*(1), 69-75.

Cotter, J. D., Patterson, M. J., & Taylor, N. A. (1995). The topography of eccrine sweating in humans during exercise. *European journal of applied physiology and occupational physiology*, *71*(6), 549-554.

Cramer, M. N., Bain, A. R., & Jay, O. (2012). Local sweating on the forehead, but not forearm, is influenced by aerobic fitness independently of heat balance requirements during exercise. *Experimental physiology*, *97*(5), 572-582.

Fealey, R. D., Low, P. A., & Thomas, J. E. (1989, June). Thermoregulatory sweating abnormalities in diabetes mellitus. In *Mayo Clinic Proceedings* (Vol. 64, No. 6, pp. 617-628). Elsevier.

Frascarolo, P., Schutz, Y., & Jequier, E. (1992). Influence of the menstrual cycle on the sweating response measured by direct calorimetry in women exposed to warm

environmental conditions. *European journal of applied physiology and occupational physiology*, *64*(5), 449-454.

Frei, R., Notley, S. R., Taylor, E. A., Burdon, C. A., Ohnishi, N., & Taylor, N. A. (2019). Revisiting the dermatomal recruitment of, and pressure-dependent influences on, human eccrine sweating. *Journal of thermal biology*, *82*, 52-62.

Fusco, M. M., Hardy, J. D., & Hammel, H. T. (1961). Interaction of central and peripheral factors in physiological temperature regulation. *American Journal of Physiology-Legacy Content*, *200*(3), 572-580.

Gagge, A. P., & Gonzalez, R. R. (2010). Mechanisms of heat exchange: biophysics and physiology. *Comprehensive physiology*, 45-84.

Gagnon, D., & Crandall, C. G. (2018). Sweating as a heat loss thermoeffector. In *Handbook of clinical neurology* (Vol. 156, pp. 211-232). Elsevier.

Gerking S and Robinson S (1946). Decline in the rates of sweating of men working in severe heat. *American Journal of Physiology* 147: 370–378.

Gisolfi CV and Wenger CB (1984) Temperature regulation during exercise: old concepts, new ideas. *Exercise and Sport Science Reviews* **12**: 339–372.

Graichen, H., Rascati, R., & Gonzalez, R. R. (1982). Automatic dew-point temperature sensor. *Journal of Applied Physiology*, *52*(6), 1658-1660.

Greenleaf, J. E., & Castle, B. L. (1971). Exercise temperature regulation in man during hypohydration and hyperhydration. *Journal of Applied Physiology*, *30*(6), 847-853.

Hammel HT. (1968). Regulation of internal body temperature. *Annual review of physiology* **30,** 641-710.

Hardy JD. (1961). Physiology of temperature regulation. *Physiological Reviews* **41,** 521-606.

Hardy JD & DuBois EF (1938). Basal metabolism, radiation, convection and vaporization at temperatures of 22 to 35ºC. J Nutr 15, 477-497.

Haslag, S. W. M., & Hertzman, A. B. (1965). Temperature regulation in young women. *Journal of Applied Physiology*, *20*(6), 1283-1288.

Hensel H. (1981). Thermoreception and temperature regulation. *Monographs of the Physiological Society* 38, 1-321.

Hertzman, A. B., Randall, W. C., Peiss, C. N., & Seckendorf, R. (1952). Regional rates of evaporation from the skin at various environmental temperatures. *Journal of applied physiology*, *5*(4), 153-161.

Hoeldtke, R. D., Bryner, K. D., Horvath, G. G., Phares, R. W., Broy, L. F., & Hobbs, G. R. (2001). Redistribution of sudomotor responses is an early sign of sympathetic dysfunction in type 1 diabetes. *Diabetes*, *50*(2), 436-443.

Hoeldtke, R. D., Bryner, K. D., & VanDyke, K. (2011). Oxidative stress and autonomic nerve function in early type 1 diabetes. *Clinical Autonomic Research*, *21*(1), 19-28.

Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports medicine*, *30*(1), 1-15.

Hopkins, W. G. (2015). Spreadsheets for Analysis of Validity and Reliability. Sportscience 19, 36-42.

Iellamo, F., Legramante, J. M., Raimondi, G., Castrucci, F., Massaro, M., & Peruzzi, G. (1996). Evaluation of reproducibility of spontaneous baroreflex sensitivity at rest and during laboratory tests. *Journal of hypertension*, *14*(9), 1099-1104.

Inoue, Nakao, Araki, & Murakami, (1991). Regional differences in the sweating responses of older and younger men. *Journal of applied physiology*, *71*(6), 2453-2459.

Inoue, Y., & Shibasaki, M. (1996). Regional differences in age-related decrements of the cutaneous vascular and sweating responses to passive heating. *European journal of applied physiology and occupational physiology*, *74*(1-2), 78-84.

Jaeschke, R., Singer, J., & Guyatt, G. H. (1989). Measurement of health status: ascertaining the minimal clinically important difference. *Controlled clinical trials*, *10*(4), 407-415.

Jay O & Kenny GP. (2010). Heat exposure in the Canadian workplace. *American Journal Industrial Medicine* **53,** 842- 853.

Kenefick, R. W., Cheuvront, S. N., Elliott, L. D., Ely, B. R., & Sawka, M. N. (2011). Biological and analytical variation of the human sweating response: implications for study design and analysis. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *302*(2), R252-R258.

Kenney, W. L., & Fowler, S. R. (1988). Methylcholine-activated eccrine sweat gland density and output as a function of age. *Journal of Applied Physiology*, *65*(3), 1082-1086.

Kenny, G. P. (2014). Muscle temperature and sweating during exercise: a new link?. *Acta Physiologica*, *212*(1), 11-13.

Kenny GP, Yardley J, Brown C, Sigal RJ & Jay O. (2010). Heat strain in older individuals and patients with common chronic diseases. *Canadian Medical Association Journal* **182,** 1053-1060.

Kenny, G. P., & Flouris, A. D. (2014). The human thermoregulatory system and its response to thermal strain. In *Protective Clothing* (pp. 319-365). Woodhead Publishing.

Kenny, G. P., & Journeay, W. S. (2010). Human thermoregulation: Separating thermal and nonthermal effects on heat loss. *Frontiers in Bioscience, 15*(25990.2)

Kenny, G. P., Yardley, J. E., Martineau, L., & Jay, O. (2008). Physical work capacity in older adults: implications for the aging worker. *American journal of industrial medicine*, *51*(8), 610-625.

Kondo N, Shibasaki M, Aoki K, Koga S, Inoue Y, and Crandall CG (2001). Function of human eccrine sweat glands during dynamic exercise and passive heat strain. *Journal of Applied Physiology* 90: 1877–1881.

Kondo, N., Takano, S., Aoki, K., Shibasaki, M., Tominaga, H., & Inoue, Y. (1998). Regional differences in the effect of exercise intensity on thermoregulatory sweating and cutaneous vasodilation. *Acta physiologica Scandinavica*, *164*(1), 71-78.

Kondo, N., Tominaga, H., Shibasaki, M., Aoki, K., Okada, S., & Nishiyasu, T. (2000). Effects of exercise intensity on the sweating response to a sustained static exercise. Journal of Applied Physiology, 88(5), 1590-1596.

Kuno, Y. (1938). Variations in secretory activity of human sweat glands. *The Lancet*, *231*(5971), 299-303.

Kuno, Y. 1956. Human perspiration. *Springfield, Illinois: Charles C Thomas publisher*.

Luo, R., Chao, C., Hsieh, P. C., Lue, J. H., & Hsieh, S. T. (2012). Effect of glycemic control on sudomotor denervation in type 2 diabetes. *Diabetes Care*, *35*(3), 612-616.

Machado-Moreira, C. A., Caldwell, J. N., Mekjavic, I. B., & Taylor, N. A. (2008a). Sweat secretion from palmar and dorsal surfaces of the hands during passive and active heating. *Aviation, space, and environmental medicine*, *79*(11), 1034-1040.

Machado-Moreira, C., Caldwell, J., van den Heuvel, A., Kerry, P., Peoples, G. E., & Taylor, N. A. (2009). Sweating and skin blood flow changes during progressive dehydration.

Machado-Moreira, C. A., Smith, F. M., van den Heuvel, A. M., Mekjavic, I. B., & Taylor, N. A. (2008b). Sweat secretion from the torso during passively-induced and exercise-related hyperthermia. *European journal of applied physiology*, *104*(2), 265-270.

Machado-Moreira, C. A., & Taylor, N. A. S. (2012a). Sudomotor responses from glabrous and non-glabrous skin during cognitive and painful stimulations following passive heating. *Acta Physiologica*, *204*(4), 571-581.

Machado-Moreira, C. A., & Taylor, N. A. (2012b). Psychological sweating from glabrous and nonglabrous skin surfaces under thermoneutral conditions. *Psychophysiology*, *49*(3), 369-374.

Machado-Moreira, C. A., Wilmink, F., Meijer, A., Mekjavic, I. B., & Taylor, N. A. (2008c). Local differences in sweat secretion from the head during rest and exercise in the heat. *European journal of applied physiology*, *104*(2), 257-264.

Matheson, G. J. (2019). We need to talk about reliability: making better use of test-retest studies for study design and interpretation. *PeerJ*, *7*, e6918.

McKinley, M. J., & Johnson, A. K. (2004). The physiological regulation of thirst and fluid intake. *Physiology*, *19*(1), 1-6.

McKinley, M. J., Martelli, D., Pennington, G. L., Trevaks, D., & McAllen, R. M. (2018). Integrating competing demands of osmoregulatory and thermoregulatory homeostasis. *Physiology, 33*(3), 170-181.

Meade, R. D., Louie, J. C., Poirier, M. P., McGinn, R., Fujii, N., & Kenny, G. P. (2016). Exploring the mechanisms underpinning sweating: the development of a specialized ventilated capsule for use with intradermal microdialysis. *Physiological reports*, *4*(6), e12738.

Morrison, S. F., & Nakamura, K. (2011). Central neural pathways for thermoregulation. *Frontiers in bioscience: a journal and virtual library*, *16*, 74.

Notley, S. R., Dervis, S., Poirier, M. P., & Kenny, G. P. (2019). Menstrual cycle phase does not modulate whole body heat loss during exercise in hot, dry conditions. *Journal of Applied Physiology*, *126*(2), 286-293.

Patterson, M.J., Stocks, J.M., and Taylor, N.A.S. (2004). Humid heat acclimation does not elicit a preferential sweat redistribution towards the limbs. *American Journal of Physiology.* 286:R512-R518.

Park, S. J., & Tamura, T. (1992). Distribution of evaporation rate on human body surface. *The Annals of physiological anthropology*, *11*(6), 593-609.

Petrofsky, J., Berk, L., & Al-Nakhli, H. (2012). The influence of autonomic dysfunction associated with aging and type 2 diabetes on daily life activities. *Experimental diabetes research*, *2012*.

Petrofsky, J., Lee, S., Patterson, C., Cole, M., & Stewart, B. (2005). Sweat production during global heating and during isometric exercise in people with diabetes. *Medical Science Monitor*, *11*(11), CR515-CR521.

Rawson, R. O., & Randall, W. C. (1961). Vascular and sweating responses to regional heating of the body surface. *Journal of applied physiology*, *16*(6), 1006-1010.

Relf, R., Willmott, A., Flint, M. S., Beale, L., & Maxwell, N. (2019). Reliability of a wearable sweat rate monitor and routine sweat analysis techniques under heat stress in females. *Journal of thermal biology*, *79*, 209-217.

Sato, K., & Dobson, R. L. (1970). Regional and individual variations in the function of the human eccrine sweat gland. *Journal of Investigative Dermatology*, *54*(6), 443-449.

Sato, K., & Sato, F. (1983). Individual variations in structure and function of human eccrine sweat gland. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *245*(2), R203-R208.

Sato, K., Kang, W. H., Saga, K., & Sato, K. T. (1989). Biology of sweat glands and their disorders. I. Normal sweat gland function. *Journal of the American Academy of Dermatology*, *20*(4), 537-563.

Sawka, M. N., Toner, M. M., Francesconi, R. P., & Pandolf, K. B. (1983). Hypohydration and exercise: Effects of heat acclimation, gender, and environment. *Journal of Applied Physiology, 55*(4), 1147-1153.

Shibasaki M & Crandall CG. (2010). Mechanisms and controllers of eccrine sweating in humans. *Frontiers in bioscience (Scholar edition)* **2,** 685-696.

Shibasaki, M., Aoki, K., Morimoto, K., Johnson, J. M., & Takamata, A. (2009). Plasma hyperosmolality elevates the internal temperature threshold for active thermoregulatory vasodilation during heat strain in humans. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*.

Shibasaki, M., Kondo, N., & Crandall, C. G. (2003). Non-thermoregulatory modulation of sweating in humans. *Exercise and Sport Sciences Reviews, 31*(1), 34-39.

Shibasaki, M., Wilson, T. E., & Crandall, C. G. (2006). Neural control and mechanisms of eccrine sweating during heat strain and exercise. *Journal of Applied Physiology*, *100*(5), 1692-1701.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological bulletin*, *86*(2), 420.

Smith, C. J., & Havenith, G. (2011). Body mapping of sweating patterns in male athletes in mild exercise-induced hyperthermia. *European journal of applied physiology*, *111*(7), 1391- 1404.

Smith, C. J., & Havenith, G. (2012). Body mapping of sweating patterns in athletes: a sex comparison. *Med Sci Sports Exerc*, *44*(12), 2350-2361.

Smith, C. J., & Havenith, G. (2019). Upper body sweat mapping provides evidence of relative sweat redistribution towards the periphery following hot-dry heat acclimation. *Temperature*, *6*(1), 50-65.

Smith, C. J., Alexander, L. M., & Kenney, W. L. (2013). Nonuniform, age-related decrements in regional sweating and skin blood flow. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *305*(8), R877-R885.

Smith, C. J., Kenney, W. L., & Alexander, L. M. (2013). Regional relation between skin blood flow and sweating to passive heating and local administration of acetylcholine in young, healthy humans. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *304*(7), R566-R573.

Stephenson, L. A., & Kolka, M. A. (1985). Menstrual cycle phase and time of day alter reference signal controlling arm blood flow and sweating. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, 249(2), R186-R191.

Stephenson, L. A., Kolka, M. A., & (With the Technical Assistance of Jane De Luca). (1999). Esophageal temperature threshold for sweating decreases before ovulation in premenopausal women. *Journal of applied physiology*, *86*(1), 22-28.

Stratford, P. W., Binkley, J. M., Riddle, D. L., & Guyatt, G. H. (1998). Sensitivity to change of the Roland-Morris back pain questionnaire: Part 1. *Physical therapy*, *78*(11), 1186-1196.

Takamata, A., Mack, G. W., Gillen, C. M., Jozsi, A. C., & Nadel, E. R. (1995). Osmoregulatory modulation of thermal sweating in humans: reflex effects of drinking. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *268*(2), R414-R422.

Takamata, A., Nagashima, K., Nose, H & Morimoto, T. (1997). Osmoregulatory inhibition of thermally induced cutaneous vasodilation in passively heated humans. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *273*(1), R197-R204.

Taylor, N. A. (2011). Human heat adaptation. *Comprehensive Physiology*, *4*(1), 325-365.

Taylor, N. A. (2006). Challenges to temperature regulation when working in hot environments. *Industrial Health* 44, 331-344

Taylor, N. A., Caldwell, J. N., & Mekjavic, I. B. (2006). The sweating foot: local differences in sweat secretion during exercise-induced hyperthermia. *Aviation, space, and environmental medicine*, *77*(10), 1020-1027.

Taylor, N. A., & Machado-Moreira, C. A. (2013). Regional variations in transepidermal water loss, eccrine sweat gland density, sweat secretion rates and electrolyte composition in resting and exercising humans. *Extreme physiology & medicine*, *2*(1), 4.

Tew, G. A., Klonizakis, M., Moss, J., Ruddock, A. D., Saxton, J. M., & Hodges, G. J. (2011). Reproducibility of cutaneous thermal hyperaemia assessed by laser Doppler flowmetry in young and older adults. *Microvascular research*, *81*(2), 177-182.

Todd, G., Gordon, C. J., Groeller, H., & Taylor, N. A. (2014). Does intramuscular thermal feedback modulate eccrine sweating in exercising humans?. *Acta Physiologica*, *212*(1), 86-96.

Van Beaumont, W., & Bullard, R. W. (1963). Sweating: its rapid response to muscular work. Science, 141(3581), 643-646.

Wendt, D., Van Loon, L. J., & Lichtenbelt, W. D. M. (2007). Thermoregulation during exercise in the heat. *Sports medicine*, *37*(8), 669-682.

PART SIX: APPENDICES

Appendix A: Ethics Certificate

16/01/2020

Université d'Ottawa
Bureau d'éthique et d'intégrité de la recherche

University of Ottawa
Office of Research Ethics and Integrity

CERTIFICAT D'APPROBATION ÉTHIQUE I CERTIFICATE OF ETHICS APPROVAL

Numéro du dossier / Ethics File Number	H-11-18-1186
Titre du projet / Project Title	Human thermoregulation: separating thermal and nonthermal effects on the body's capacity to dissipate heat
Type de projet / Project Type	Recherche de professeur / Professor's research project
Statut du projet / Project Status	Renouvelé / Renewed
Date d'approbation (jj/mm/aaaa) / Approval Date (dd/mm/yyyy)	16/01/2020
Date d'expiration (jj/mm/aaaa) / Expiry Date (dd/mm/yyyy)	10/01/2021

Équipe de recherche / Research Team

Chercheur / Researcher	Affiliation	Role
Glen KENNY	École des sciences de l'activité physique / School of Human Kinetics	Chercheur Principal / Principal Investigator
Kelli KING	University of Ottawa	Étudiant-chercheur / Student-researcher
Melissa CÔTÉ	University of Ottawa	Étudiant-chercheur / Student-researcher
Ronald SIGAL		Co-chercheur / Co-investigator
Madison SCHMIDT	École des sciences de l'activité physique / School of Human Kinetics	Étudiant-chercheur / Student-researcher
Caroline Mutheu MUIA	École des sciences de l'activité physique / School of Human Kinetics	Étudiant-chercheur / Student-researcher
Emma MCCOURT	Département de biologie / Department of Biology	Étudiant-chercheur / Student-researcher
Ashley AKERMAN	École des sciences de l'activité physique / School of Human Kinetics	Co-chercheur / Co-investigator
Jeremy MCCORMICK	École des sciences de l'activité physique / School of Human Kinetics	Co-chercheur / Co-investigator
Emileigh BINET	École des sciences de l'activité physique / School of Human Kinetics	Étudiant-chercheur / Student-researcher
Morgan MCMANUS	École des sciences de l'activité physique / School of Human Kinetics	Étudiant-chercheur / Student-researcher
Samah SACI	École des sciences de l'activité physique / School of Human Kinetics	Étudiant-chercheur / Student-researcher
Brian FRIESEN	École des sciences de l'activité physique / School of Human Kinetics	Étudiant-chercheur / Student-researcher
Marcel RUZICKA		Co-chercheur / Co-investigator
Maura RUTHERFORD	École des sciences de l'activité physique / School of Human Kinetics	Étudiant-chercheur / Student-researcher
Robert MEADE	École des sciences de l'activité physique / School of Human Kinetics	Étudiant-chercheur / Student-researcher
Andrew D'SOUZA	École des sciences de l'activité physique / School of Human Kinetics	Étudiant-chercheur / Student-researcher
Sean NOTLEY	École des sciences de l'activité physique / School of Human Kinetics	Co-chercheur / Co-investigator
	École des sciences de l'activité physique / School of Human	

550, rue Cumberland, pièce 154
Ottawa (Ontario) K1N 6N5 Canada

550 Cumberland Street, Room 154
Ottawa, Ontario K1N 6N5 Canada

613-562-5387 • 613-562-5338 • ethique@uOttawa.ca / ethics@uOttawa.ca
www.recherche.uottawa.ca/deontologie I www.recherche.uottawa.ca/ethics

Université d'Ottawa
Bureau d'éthique et d'intégrité de la recherche

University of Ottawa
Office of Research Ethics and Integrity

Gregory MCGARR	Kinetics	Co-chercheur / Co-investigator
Serena TOPSHEE	Département de biologie / Department of Biology	Étudiant-chercheur / Student-researcher
Mohamed GEMAE	École des sciences de l'activité physique / School of Human Kinetics	Étudiant-chercheur / Student-researcher
Pierre BOULAY		Co-chercheur / Co-investigator

Conditions spéciales ou commentaires / Special conditions or comments

The study has a sub-title that is: "**Understanding the interplay of thermal and nonthermal factors on whole-body heat exchange during heat stress**."

16/01/2020

Université d'Ottawa
Bureau d'éthique et d'intégrité de la recherche

University of Ottawa
Office of Research Ethics and Integrity

Le Comité d'éthique de la recherche (CÉR) de l'Université d'Ottawa, opérant conformément à l'*Énoncé de politique des Trois conseils* (2014) et toutes autres lois et tous règlements applicables, a examiné et approuvé la demande d'éthique du projet de recherche ci-nommé.

L'approbation est valide pour la durée indiquée plus haut et est sujette aux conditions énumérées dans la section intitulée "Conditions Spéciales ou Commentaires". Le formulaire « Renouvellement ou Fermeture de Projet » doit être complété quatre semaines avant la date d'échéance indiquée ci-haut afin de demander un renouvellement de cette approbation éthique ou afin de fermer le dossier.

Toutes modifications apportées au projet doivent être approuvées par le CÉR avant leur mise en place, sauf si le participant doit être retiré en raison d'un danger immédiat ou s'il s'agit d'un changement ayant trait à des éléments administratifs ou logistiques du projet. Les chercheurs doivent aviser le CÉR dans les plus brefs délais de tout changement pouvant augmenter le niveau de risque aux participants ou pouvant affecter considérablement le déroulement du projet, rapporter tout évènement imprévu ou indésirable et soumettre toute nouvelle information pouvant nuire à la conduite du projet ou à la sécurité des participants.

The University of Ottawa Research Ethics Board, which operates in accordance with the *Tri-Council Policy Statement* (2014) and other applicable laws and regulations, has examined and approved the ethics application for the above-named research project.

Ethics approval is valid for the period indicated above and is subject to the conditions listed in the section entitled "Special Conditions or Comments". The "Renewal/Project Closure" form must be completedfour weeks before the above-referenced expiry date to request a renewal of this ethics approval or closure of the file.

Any changes made to the project must be approved by the REB before being implemented, except when necessary to remove participants from immediate endangerment or when the modification(s) only pertain to administrative or logistical components of the project. Investigators must also promptly alert the REB of any changes that increase the risk to participant(s), any changes that considerably affect the conduct of the project, all unanticipated and harmful events that occur, and new information that may negatively affect the conduct of the project or the safety of the participant(s).

Germain ZONGO
Responsable d'éthique en recherche / Protocol Officer
Pour/For **Daniel LAGAREC** Président(e) du/ Chair of the **Comité d'éthique de la recherche en sciences de la santé et sciences / Health Sciences and Sciences Research Ethics Board**

550, rue Cumberland, pièce 154
Ottawa (Ontario) K1N 6N5 Canada

550 Cumberland Street, Room 154
Ottawa, Ontario K1N 6N5 Canada

613-562-5387 • 613-562-5338 • ethique@uOttawa.ca / ethics@uOttawa.ca
www.recherche.uottawa.ca/deontologie | www.recherche.uottawa.ca/ethics