Skin Temperature Measurement Using Contact Thermometry: A Systematic Review of Setup Variables and Their Effects on Measured Values

Braid A. MacRae[1,2], Simon Annaheim[1]*, Christina M. Spengler[2,3†] and René M. Rossi[1†]

[1] Empa, Swiss Federal Laboratories for Materials Science and Technology, Laboratory for Biomimetic Membranes and Textiles, St. Gallen, Switzerland, [2] Exercise Physiology Lab, Institute of Human Movement Sciences and Sport, Department of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland, [3] Zurich Center for Integrative Human Physiology, University of Zurich, Zurich, Switzerland

Background: Skin temperature (T_{skin}) is commonly measured using T_{skin} sensors affixed directly to the skin surface, although the influence of setup variables on the measured outcome requires clarification.

Objectives: The two distinct objectives of this systematic review were (1) to examine measurements from contact T_{skin} sensors considering equilibrium temperature and temperature disturbance, sensor attachments, pressure, environmental temperature, and sensor type, and (2) to characterise the contact T_{skin} sensors used, conditions of use, and subsequent reporting in studies investigating sports, exercise, and other physical activity.

Data sources and study selection: For the measurement comparison objective, Ovid Medline and Scopus were used (1960 to July 2016) and studies comparing contact T_{skin} sensor measurements *in vivo* or using appropriate physical models were included. For the survey of use, Ovid Medline was used (2011 to July 2016) and studies using contact temperature sensors for the measurement of human T_{skin} *in vivo* during sport, exercise, and other physical activity were included.

Study appraisal and synthesis methods: For measurement comparisons, assessments of risk of bias were made according to an adapted version of the Cochrane Collaboration's risk of bias tool. Comparisons of temperature measurements were expressed, where possible, as mean difference and 95% limits of agreement (LoA). Meta-analyses were not performed due to the lack of a common reference condition. For the survey of use, extracted information was summarised in text and tabular form.

Results: For measurement comparisons, 21 studies were included. Results from these studies indicated minor ($<0.5°C$) to practically meaningful ($>0.5°C$) measurement bias within the subgroups of attachment type, applied pressure, environmental conditions, and sensor type. The 95% LoA were often within $1.0°C$ for *in vivo* studies and $0.5°C$ for

physical models. For the survey of use, 172 studies were included. Details about T_{skin} sensor setup were often poorly reported and, from those reporting setup information, it was evident that setups widely varied in terms of type of sensors, attachments, and locations used.

Conclusions: Setup variables and conditions of use can influence the measured temperature from contact T_{skin} sensors and thus key setup variables need to be appropriately considered and consistently reported.

INTRODUCTION

Heat exchanges at the skin surface can both contribute to and challenge thermal homeostasis. Accordingly, quantification of skin temperature (T_{skin}) is important in many research and applied settings and using sensors affixed directly to the skin surface is common for this purpose. While measuring T_{skin} in this way is simple in terms of access to the measurement site, an inherent challenge is that accurate measurements of surface temperature are difficult to accomplish (Hardy, 1934b; Malone, 1963; Boetcher et al., 2009). Understanding measurement limitations benefits users by supporting decisions during sensor selection and setup, use, and interpretation of resultant data, and can assist in the development of future sensor systems.

Human T_{skin} data are used widely, including for evaluating thermal strain (Moran et al., 1999; International Organization for Standardization, 2004; McLellan et al., 2013), estimating mean body temperature (Burton, 1935; Colin et al., 1971; Lenhardt and Sessler, 2006) and body heat content and storage by thermometry (Tucker et al., 2006) (although with limitation compared to whole body calorimetry; Jay et al., 2007, 2010), and validating psychophysical and thermophysiological models (Zhang et al., 2010; Martinez et al., 2016). Further, T_{skin} is used for understanding the mechanisms of acute thermoregulatory and cardiovascular responses (Rowell et al., 1969; Nadel et al., 1971; Libert et al., 1978; Cotter and Taylor, 2005) and adaptation to thermal stress (Regan et al., 1996), understanding how perturbations of body temperature, and strategies to mitigate or recover from such perturbations, influence aspects of sports, and occupational performance (Daanen, 2009; Sawka et al., 2011, 2012; Schlader et al., 2011a; Caldwell et al., 2012; Levels et al., 2012; Ross et al., 2013; Lee et al., 2015; Stevens et al., 2017), and understanding effects of the ambient environment (Gagge et al., 1967; Galloway and Maughan, 1997) or body coverings (Gagge et al., 1938; Nielsen et al., 1985; White and Hodous, 1987; Ruckman et al., 1999; Rossi, 2003; Barwood et al., 2016). With growing interest in wearable technology and body monitoring, the prospects for applied uses of T_{skin} data are also expanding and include monitoring of individuals during sport and exercise or working under extreme conditions.

Contact thermometry consists of a temperature sensor positioned in direct contact with the skin surface and, therefore, relies on conductive heat exchanges between skin and sensor. Contact thermometry is popular for measuring T_{skin} in both research and applied settings with reasons including commercial sensor availability and relative low cost, small sensor size and potential robustness, ease of measurement continuity within a measurement period, the ability to position sensors at a selection of body sites, and that measurements can be made at sites under coverings like clothing and protective equipment. These sensors, however, can differ in terms of the underlying measurement principle (e.g., electrical resistance for thermistors, Seebeck effect for thermocouples), sensor size and shape, constituent materials, and resultant thermal properties. The method of attachment to the skin also varies, with tape or other adhesive common in research settings (Buono and Ulrich, 1998; Psikuta et al., 2014). With the variety of contact T_{skin} sensors and attachments commonly used, it is necessary to recognise if equivalent measurements can be expected irrespective of the sensor type and setup variables.

Issues regarding validity of data from contact T_{skin} sensors also require consideration. That is, do these sensors actually measure what they are supposed to be measuring? Due to inherent skin contact and coverage at the measured sites, a disadvantage of contact thermometry is that the sensors and their attachments modify the immediate environment for the underlying skin. The effects of such a modification may manifest as a disturbance of the temperature that would otherwise exist in the undisturbed case (Henriques, 1947; Childs et al., 2000; Boetcher et al., 2009; Taylor et al., 2014). Further, the T_{skin} is not measured directly: the temperature measured is that of the sensor itself (Malone, 1963; Childs, 2001) and, therefore, it is important to consider whether the sensor temperature is at equilibrium with the underlying skin temperature. Non-contact techniques utilising electromagnetic radiation emitted from the skin surface—therefore, able to measure temperature from a distance—can overcome certain difficulties associated with sensor contact and coverage (Hardy, 1934a,b; Hardy and Soderstrom, 1937), but have difficulties of their own, including limited practicality during movement or field use, influences of changes in the emissivity of the skin surface (Bernard et al., 2013), and unclear validity of common commercial devices, particularly during exercise (Bach et al., 2015b). Thus, no method is without limitation and, while imperfect, the general practicality of contact thermometry likely means it will remain a desirable method for quantifying T_{skin} in the foreseeable future.

To elucidate relevant considerations for the measurement of T_{skin} using contact thermometry, we systematically reviewed the literature in accordance with two distinct objectives. The primary objective, objective 1, was to examine measurements from

contact T_{skin} sensors considering sensor-surface equilibrium temperature and surface temperature disturbance, and considering whether setup variables (sensor type, attachments, applied pressure, or environmental conditions) meaningfully influence the measured temperature. The secondary objective, objective 2, was to survey and characterise the contact T_{skin} sensors used, conditions of use, and subsequent reporting of T_{skin} in published scientific studies investigating sport, exercise, and other physical activity.

METHODS

A protocol was developed before the formal searches were performed, specifying the literature searches, screening, inclusion, and data synthesis. The review was conducted in general accordance with PRISMA guidelines (Moher et al., 2009).

Search Strategy

The search strategies (including full lists of search terms) used for both objectives are given in Supplementary Material Tables 1, 2. In brief, terms including and relating to skin, temperature, and thermometry were used for the measurement comparisons (objective 1), and terms including and relating to skin, temperature, sport, exercise, and physical activity were used for the survey of use (objective 2). Searches were performed 13.07.2016.

The electronic databases Ovid Medline and Scopus were used for objective 1. Reference lists of studies meeting inclusion criteria were searched for other relevant articles and Google Scholar was used to identify any relevant articles citing included studies. The electronic database Ovid Medline was used for objective 2. One database was considered appropriate because, given the purpose of this objective and the breadth of journals included in Medline, the practicality was considered to outweigh the risk of bias for publication sampling. No further articles were added for objective 2 from other search means. No language restrictions were imposed for the searches. Studies published in the grey literature or in truncated form (e.g., abstracts only, posters) were not considered.

Study Eligibility Criteria
Measurement Comparisons

We sought temperature data from contact temperature sensors. Studies using "conventional" contact temperature sensors—those that can be affixed to the skin surface to measure T_{skin}—were included; other contact sensors (including those temporarily held against the skin) were included only in cases where they provided insight into setup considerations for affixable sensors. Measurements from liquid-crystal thermometers or other colour-change materials and measurements from non-contact methods were excluded. Studies were included if human or other animal T_{skin} was measured *in vivo*; surface temperature ($T_{surface}$) measurements of a physical model were also included if the model was used in the context of understanding human T_{skin} measurement and if the temperatures investigated were physiologically relevant for humans. Included studies

were required to report the measurement of a suitable comparison temperature, typically another T_{skin} or $T_{surface}$ (the comparator determined to what outcome subgroup the data belonged; see data synthesis below). Comparisons performed simply for the purpose of validating or correcting a prototype sensor system were excluded (e.g., Chen et al., 2015). Comparisons with non-contact methods or numerical models, while useful in their own right, were beyond the scope of this article. (The comparison of contact and infra-red devices for measuring T_{skin} has been reviewed elsewhere; Bach et al., 2015b.) Studies investigating T_{skin} or $T_{surface}$ during clinical or rehabilitation treatments/conditions were excluded (e.g., cryotherapy, ultrasound therapy, radiotherapy, other hyperthermia therapy). Studies published from 1960 (until 13.07.2016) were considered for inclusion, which was a modification from the unrestricted date range specified in the protocol (see Supplementary Material Appendix 1).

Survey of Use

We sought general information about T_{skin} sensors and their use in published research (e.g., sensor type, attachments, conditions of sensor use, and use of the T_{skin} data). Studies included were those that used contact temperature sensors for the measurement of human T_{skin} *in vivo* during experimental studies involving sport, exercise, and other physical activity (PA; hereafter for practicality, sport, exercise, occupational and other PA will be collectively termed PA). The sensors had to be affixed to the skin surface in some way and remain in place during the PA. While measurement of T_{skin} is relevant beyond situations involving PA (e.g., clinical use, passive heat stress, sleep studies, circadian rhythm studies), measurements in these contexts were beyond the scope of this review.

For inclusion, studies needed to report T_{skin} data recorded during or immediately following PA, or use that data to calculate another variable that was reported in the results. Studies were excluded if they involved only "trivial" PA, such as standing tasks. Studies in which it was clear the T_{skin} data had been previously published in article form and studies involving daily monitoring were excluded unless that monitoring was part of a specific occupational routine including an expected component of PA. For objective 2, only studies published in the years 2011–2016 (until 13.07.2016) were included to be reflective of current practice at the time of review.

Study Selection and Data Extraction

All records were screened against the eligibility criteria first by title and abstract then, for those remaining, by full text. For measurement comparisons, duplicates were removed and preliminary inclusion of studies was performed by one investigator (BM). Final inclusion was performed after a systematic cross-check (SA or CS); disagreements were resolved by discussion among the three investigators. For the survey of use, one investigator (BM) performed preliminary inclusion. Final inclusion was performed after a systematic cross-check (SA) and any disagreements were resolved by discussion with a third investigator (CS).

For measurement comparisons, one investigator (BM) extracted study information and temperature data, with a second investigator (SA) verifying the accuracy. Disagreements were solved by discussion including a third investigator (CS). Temperature data were typically extracted from text or tables; for eight studies it was necessary to extract data, at least in part, from figures (Flesch et al., 1976; Dollberg et al., 1994; Buono and Ulrich, 1998; van Marken Lichtenbelt et al., 2006; Deng and Liu, 2008; Youhui et al., 2010; McFarlin et al., 2015; Priego Quesada et al., 2015) and original data were available for two studies (Psikuta et al., 2014; Bach et al., 2015a). Data extracted from figures was done so using a computer-based extraction tool (http://arohatgi.info/WebPlotDigitizer/). For the survey of use, one investigator extracted study information (BM) and at least one other investigator was consulted to discuss uncertainties. Spreadsheets were produced and refined during piloting work with input from multiple investigators and common problems were discussed before formal data extraction. Further details about the extracted data information are given in Supplementary Material Appendix 2.

Assessments of Risk of Bias

For measurement comparisons, assessments of risk of bias were made at the outcome subset level according to an adapted version of the Cochrane Collaboration's risk of bias tool (Higgins et al., 2011) (see Supplementary Material Table 3 for the specific criteria used). Seven sources of bias were assessed: four that are widely used (sequence generation, blinding of participants and personnel, incomplete outcome data, selective reporting) and three specific "other" sources of bias (consistency of test conditions, calibration/baseline comparability of sensors, and study support). Assessments were made independently by two investigators (BM, SA); a third investigator was available for mediation but was not required. Sources of bias were judged as low, unclear, or high risk of material bias (Supplementary Material Table 3). Insufficient detail to make an assessment was considered an unclear risk of bias. Information that was reported but with an unclear influence on bias was also classified as unclear risk of bias. Risk of bias was assessed from the perspective of this review, which may differ from the perspective of the original study. Assessments of risk of bias were not applicable for the survey of use (objective 2) due to scope.

Data Analysis and Synthesis

For the survey of use, the collected information was compiled in Microsoft® Excel and summarised in text and table form. The remaining information on data analysis and synthesis below applies to the measurement comparisons.

The identified temperature measurement comparisons were categorised according to six pre-defined concept-based outcome subgroups for presentation and synthesis (1. temperature disturbance of the surface underlying a surface sensor, 2. thermal equilibrium of the surface sensor with the underlying temperature, 3. influence of the attachment on surface sensors, 4. influence of the pressure applied by surface sensors, 5. influence of the environmental conditions on surface sensors, 6. influence of the type of surface sensor; descriptions in

Supplementary Material Appendix 3). Specific details about comparisons differed from study to study (e.g., number of sensors and sites used, timing of comparison measurements) and therefore such information was retained.

Comparisons were expressed as mean temperature differences (estimate of measurement bias; calculated as comparator 2–comparator 1) and 95% limits of agreement (LoA; estimate of random error), in °C. The 95% LoA were calculated, where possible, as:

$$95\% \text{ LoA} = \text{mean difference } \pm t_{n-1} \cdot s_{diff} \qquad (1)$$

where t_{n-1} is the corresponding critical value from the t-distribution, n is the sample size, and s_{diff} is the standard deviation of the differences. Due to being typically limited to summary data, heteroscedasticity and normality of the individual differences were assumed and not directly assessed. We used the critical value from the t-distribution for the calculation of LoA (cf. 1.96 or 2; Bland and Altman, 1986) because the sample sizes here were commonly <15 (Williamson et al., 2002). There were cases in which the s_{diff} was not available. In cases without any form of variance presented, the point estimates were used without LoA. In cases where other suitable parameters were reported (e.g., confidence interval for the mean difference, the standard deviation of each comparator separately), we were able to estimate s_{diff} (for further detail see Supplementary Material Appendix 4).

Studies that were unable to be presented in figures were acknowledged in narrative form (a lack of information in the original study was the typical reason). For practicality reasons, for cases in which there were more than two comparators within a specific study outcome (e.g., a comparison of seven levels of applied pressure; Jirak et al., 1975), one was designated as the "common comparator" from which the others were compared (in this example, at five body sites separately, the measured temperature with an applied pressure of 136 mmHg was compared to the measured temperature using, in series, the same sensor at the same body site with applied pressures of 34–681 mmHg). The common comparator was either dictated by the data presentation in the original article (e.g., data from Jirak et al., 1975) or it was selected by the review investigators after considering the original context of the data and the purpose of the comparison in the context of this review (e.g., data from Psikuta et al., 2014).

The interpretation of mean differences was made cognizant of the lack of a universal reference T_{skin} or $T_{surface}$. Thus, comparisons were typically relative, without one measurement being considered more accurate (a better estimate of the so-called "true" value; an exception here is the temperature disturbance and thermal equilibrium subgroups in which the reference temperatures were assumed to be superior to the surface T_{skin} sensor). Similarly, the LoA were not considered to be a perfect or complete indicator of random error, but rather a reasonable and familiar estimate. The threshold for being considered practically meaningful was beyond ±0.5°C for mean difference (*in vivo* and physical models) and, for LoA, ±1.0°C for *in vivo* Tskin (similar to those used elsewhere; Harper Smith et al., 2010; James

et al., 2014) and $\pm 0.5°C$ for physical models. These thresholds were a simplification of reality but necessary for practicality. The thresholds for human T_{skin} were supported by retest reliability data from contact temperature sensors (thermistors; for mean T_{skin} during rest and exercise in sessions separated by 1 week, mean differences were $<0.5°C$ and LoA typically $<1.0°C$) (James et al., 2014). The physiological relevance of these thresholds will, in practice, depend upon a particular application (e.g., purpose of the measurement, number of sites used, the requirement for external validity of absolute values). Lower LoA thresholds were used for the physical models assuming greater experimental control of a model versus human skin. Meta-analyses were not performed due to inconsistency across studies, particularly in terms of reference condition (comparator sensor type, setup, and experimental conditions).

Data were displayed, where possible, in forest plots (Supplementary Material) and condensed into summary plots (main text). Summary plots were used for initial data display to facilitate an accessible overview while the forest plots were given in the Supplementary Material to retain the option for a more detailed examination of specific comparisons. Data for summary plots (mean, minimum, and maximum for mean differences and LoA) were calculated from the data presented in the forest plots with the mean differences first converted into absolute values (i.e., mean absolute error). Therefore, mean differences in the summary plots indicate only the magnitude of the difference, whereas mean differences in the forest plots indicate both the magnitude and relative direction.

RESULTS

For the measurement comparisons objective, study characteristics, assessment of risk of bias, and outcomes are presented in separate sections below. For the survey of use objective, the outcomes of interest were simply study characteristics and, therefore, all results are contained within section Study Characteristics.

Study Characteristics
An overview of the literature search and study screening for the two objectives is given in **Figure 1**.

Measurement Comparisons
The search yielded 4,299 records (**Figure 1**). Following removal of duplicates ($n = 1,095$) and exclusion during screening ($n = 3,184$), 20 studies remained. An additional study was located from searches of the reference lists. Thus, 21 studies were finally included (Yakovlev and Utekhin, 1965, 1966; Guadagni et al., 1972; Jirak et al., 1975; Flesch et al., 1976; Mahanty and Roemer, 1979a,b; Krause, 1993; Dollberg et al., 1994; Lee et al., 1994; Buono and Ulrich, 1998; van Marken Lichtenbelt et al., 2006; Deng and Liu, 2008; Harper Smith et al., 2010; Youhui et al., 2010; Tyler, 2011; James et al., 2014; Psikuta et al., 2014; Bach et al., 2015a; McFarlin et al., 2015; Priego Quesada et al., 2015). From these 21 studies, 31 distinct subsets of comparisons were identified. A subset is considered here as data from a particular study that addresses a distinc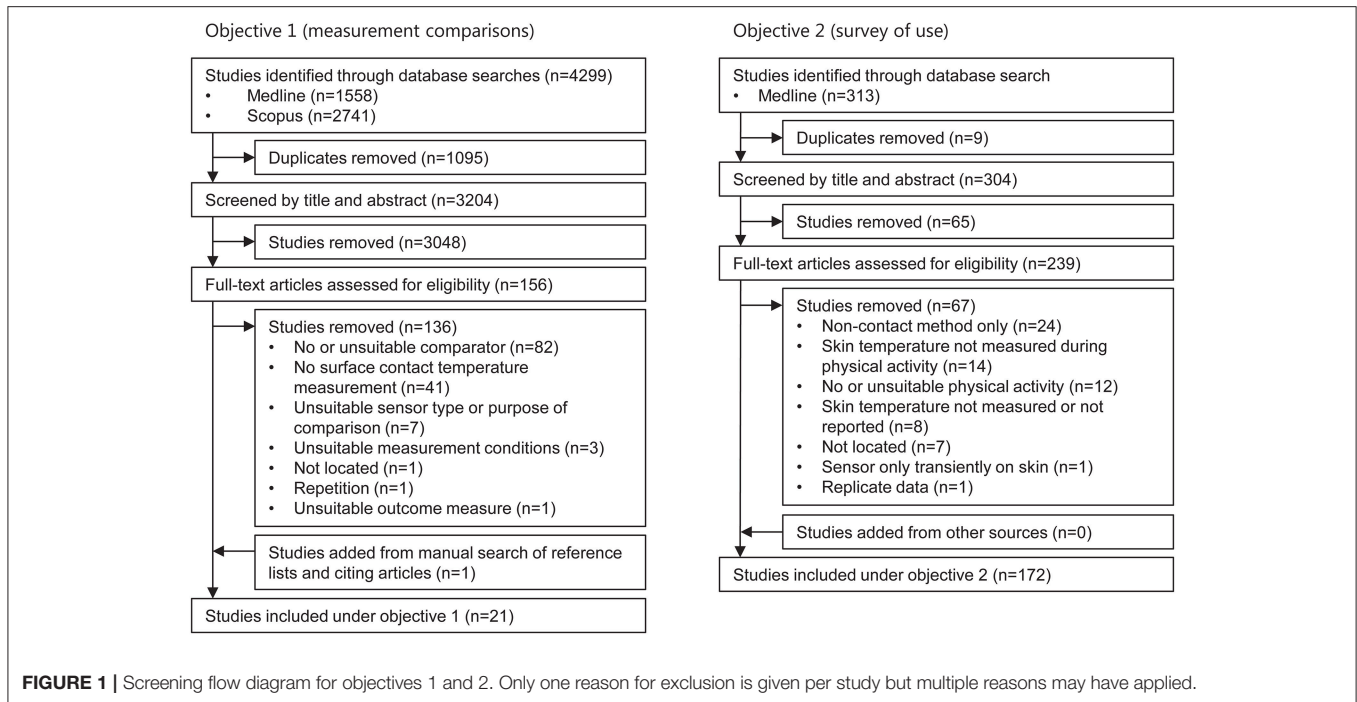t aspect of one (or more) of the outcome subgroups. Four subsets (Buono and Ulrich, 1998; van Marken Lichtenbelt et al., 2006; Harper Smith et al., 2010; Tyler, 2011) were applicable to two outcome subgroups and one study (Psikuta et al., 2014) was applicable to four, giving 38 subsets in total.

Selected general information from the included studies is given in Supplementary Material Table 4. From the 21 studies included, the relevant outcome data comprised measurements from a physical model in nine studies (Yakovlev and Utekhin, 1965, 1966; Flesch et al., 1976; Mahanty and Roemer, 1979a; Krause, 1993; Lee et al., 1994; James et al., 2014; Psikuta et al., 2014; Priego Quesada et al., 2015) and measurements from the skin of human subjects in 17 studies (Yakovlev and Utekhin, 1965, 1966; Guadagni et al., 1972; Jirak et al., 1975; Mahanty and Roemer, 1979a,b; Dollberg et al., 1994; Lee et al., 1994; Buono and Ulrich, 1998; van Marken Lichtenbelt et al., 2006; Deng and Liu, 2008; Harper Smith et al., 2010; Youhui et al., 2010; Tyler, 2011; James et al., 2014; Bach et al., 2015a; McFarlin et al., 2015). In human studies, T_{skin} data was reported from single body sites in 10 studies (Yakovlev and Utekhin, 1965, 1966; Jirak et al., 1975; Mahanty and Roemer, 1979a,b; Dollberg et al., 1994; Lee et al., 1994; Deng and Liu, 2008; Youhui et al., 2010; McFarlin et al., 2015) and as the mean of multiple sites in seven studies (Buono and Ulrich, 1998; van Marken Lichtenbelt et al., 2006; Harper Smith et al., 2010; Youhui et al., 2010; Tyler, 2011; James et al., 2014; Bach et al., 2015a); in two studies it was unclear if the data represented single sites or a mean of multiple sites (Yakovlev and Utekhin, 1966; Mahanty and Roemer, 1979b). The range for the number of skin sites used within a study was 1–14 (3–14 for those that reported mean T_{skin}).

Survey of Use
This second search yielded 313 records (**Figure 1**). Following removal of duplicates ($n = 9$) and exclusion during screening ($n = 132$), 172 studies were retained (see in-text Appendix). These studies included 2,267 participants (2014 male, 213 female, 40 not specified). T_{skin} data were reported for participants cycling in 84 studies (49% of included studies), running in 40 studies (23%), walking in 40 studies (23%), occupational PA in 7 studies (4%), and other PA in 18 studies (11%). Most studies (90%) also reported resting T_{skin} data.

A summary of the information about the sensors, attachments, and data use is given in **Table 1**. The type of contact sensor used (e.g., thermistor, thermocouple) was reported in 84% of studies. Of the studies reporting sensor type, thermistors were the most common (89 studies) followed by thermocouples (30 studies) and iButtons (oscillator-based digital thermometer; 26 studies). Identifying the sensor manufacturer (e.g., Grant Instruments Ltd, Cambridge, UK) and model (e.g., EUS-UU-VL3-0) was not always possible: in 34 studies (20%) no manufacturer or supplier information was reported and in 65 studies (38%) no sensor model information was reported. Details about sensor calibration were rarely reported with 94% of studies providing no or unclear information. Similarly, most studies (89%) reported no error-related information about the sensor (e.g., uncertainty, precision). Over half of the studies reported no information

FIGURE 1 | Screening flow diagram for objectives 1 and 2. Only one reason for exclusion is given per study but multiple reasons may have applied.

about the method of sensor attachment: of the 73 studies (42%) reporting clear information, 63 studies used tape. It was possible to determine whether the sensors were covered or uncovered in 40% of studies, with the remaining 60% unclear or not reported. Most studies reported mean T_{skin} (83%) and almost all studies reported absolute temperature values (97%). Irrespective of the number of measurement sites used, 57 studies (33%) reported some data from single measurement sites or single sensor.

The number and location of body sites used are given in Supplementary Material Tables 5 and 6, respectively. The number of body sites per participant ranged from one to sixteen, with four sites most commonly used (46%). Chest (74%), anterior thigh (71%), lower leg (70.3%), and upper arm (56%) were the most common sites used.

Risk of Bias within Studies—Measurement Comparisons

Risk of bias for measurement comparisons was assessed at the subset level within each outcome subgroup. These assessments (given in full in Supplementary Material Figure 1) are summarised in **Figure 2**. In total, the risk of bias was typically unclear (72%), with only a quarter (24%) judged as low risk of bias; the remaining 5% were judged as high risk of bias. The domains most frequently judged as high risk of bias were calibration/baseline comparability of the sensors (11%) and selective reporting (8%). Making the assessment was often challenging due to limited reporting in the original articles: this limitation contributed to the high proportion judged as unclear risk of bias.

Outcomes—Measurement Comparisons

The ways in which individual studies or experiments were conducted and reported were diverse and inconsistent across studies (e.g., model surfaces or skin sites used, types of sensors and attachments, environments and procedures; pooled mean or individual measurement sites or timings). Therefore, meta-analyses were not performed.

The included data are presented in summary plots (**Figures 3–8**) and corresponding forest plots (Supplementary Material Figures 2–7). Data from six subsets from five studies (Yakovlev and Utekhin, 1965, 1966; Guadagni et al., 1972; Mahanty and Roemer, 1979a,b) were unable to be presented visually due to a lack of information and, therefore, are summarised briefly in the text. The LoA were not included in seven subsets from six studies due to unreported variance estimates (Yakovlev and Utekhin, 1966; Jirak et al., 1975; Flesch et al., 1976; Mahanty and Roemer, 1979b; Deng and Liu, 2008; Youhui et al., 2010).

Temperature Disturbance of the Surface Underlying a Surface Sensor

Data from one study were identified (Mahanty and Roemer, 1979a) ($n = 1$ subset, giving 2 comparisons; **Figure 3**). The surface temperature of a physical model was estimated from a thermocouple 0.4 mm below the surface, without and with a surface temperature sensor directly above (undisturbed and disturbed states, respectively). The undisturbed surface temperature was similar to both the disturbed temperature and the surface thermistor probe acting as the disturbance (mean differences of $<0.1°C$ for the point estimates).

TABLE 1 | Information from survey of sensor usage in studies involving physical activity (n = 172 studies); data are study count with percentage in parentheses.

	Reported	Unclear or not reported
Type of contact temperature sensor	144 (84%)[a]	28 (16%)
Sensor calibration	10 (6%)	162 (94%)
Sensor attachment method	73 (42%)[b]	99 (58%)
	Something reported	**Not reported**
Sensor accuracy, uncertainty, precision, etc.	19 (11%)	153 (89%)
	Covered	**Uncovered**
Sensor coverage by attachment[c]	65 (38%)	3 (2%)
	Yes	**No or unclear**
Calculations using skin temperature data		
Mean skin temperature	142 (83%)	30 (17%)
Mean body temperature	39 (23%)	133 (77%)
Other calculations	56 (33%)	116 (67%)
	Absolute	**Change score**
Skin temperature data presentation	166 (97%)	30 (17%)

[a]*Sensor types reported were: thermistors [n = 89; the most common manufacturer reported was Grant Instruments Ltd., Cambridge, UK (n = 29) followed by YSI Inc., Yellow Springs, OH, USA (n = 18); 10 studies reported no manufacturer or supplier information], thermocouples [n = 30; the most common manufacturer or supplier reported was Concept Engineering, Old Saybrook, CT, USA (n = 5) followed by Omega Engineering Ltd, Stamford, CT, USA (n = 3); 17 studies reported no manufacturer or supplier information], iButtons (an oscillator-based digital thermometer; n = 26; Maxim Integrated Products, CA, USA), and resistance thermometers (n = 2). Three studies each used two types of contact temperature sensors so the total of sensor types here is 147 (cf. 144 in the table above).*

[b]*Of the 73 reported, 63 studies used tape and 33 of those studies specified the tape type: 3 M Transpore (n = 13), 3 M Tegaderm (n = 9), BSN Medical Fixomull (n = 4), 3 M Blenderm (n = 2), 3 M Medipore (n = 2), BSN Medical Hypafix (n = 2), Hy-Tape international Hy-Tape (n = 2), Leuko Sportstape (n = 1).*

[c]*The remaining (n=104, 60%) were unclear or not reported.*

Thermal Equilibrium of the Surface Sensor with the Underlying Temperature

Data from three studies were identified (Mahanty and Roemer, 1979a; Lee et al., 1994; Psikuta et al., 2014) (n = 4 subsets, giving 96 comparisons; **Figure 4** and Supplementary Material Figure 3). Temperature of the externally applied surface sensors was typically lower than invasive T_{skin} *in vivo* (Lee et al., 1994) and lower than the model reference temperature (Lee et al., 1994; Psikuta et al., 2014), although the magnitude of the mean differences was variable, ranging from <0.5°C (Mahanty and Roemer, 1979a; Lee et al., 1994; Psikuta et al., 2014) to 5.8°C (Psikuta et al., 2014). For surface T_{skin} sensors and attachments commonly used in non-clinical human research studies (Psikuta et al., 2014), greater differences were observed

with a greater surface-environment gradient (see also influence of the environmental conditions). The data from Krause (1993) was not included here because the reference temperature sensors within the physical model are not described and no plate temperature data is reported.

Influence of the Attachment on Surface Sensors

Data from six studies were identified (Dollberg et al., 1994; Buono and Ulrich, 1998; Deng and Liu, 2008; Tyler, 2011; Psikuta et al., 2014; Priego Quesada et al., 2015) (n = 6 subsets, giving 19 comparisons; **Figure 5** and Supplementary Material Figure 4). The type of attachment used had mixed effects on the mean differences of T_{skin} or $T_{surface}$ with absolute mean differences ranging from 0.1 to 1.4°C. The LoA were typically within ±1.0°C for measurements on human skin, while those on physical models ranged from 0.3 to 1.5°C.

Using an attachment resulted in numerically greater measured temperatures by 0.1–1.3°C compared to the same sensors "uncovered" for human skin (Buono and Ulrich, 1998; Tyler, 2011) or a model surface (Priego Quesada et al., 2015). With increasing environmental temperature, the differences of covered versus "uncovered" sensors became smaller in one study (mean difference of 1.3°C in a 23°C environment, decreasing to 0.4°C in 42°C) (Buono and Ulrich, 1998) but tended to become larger in another study (mean difference of 0.1°C increasing to 0.5°C for one layer of tape, 0.8°C increasing to 0.9°C for two layers of tape plus bandage, environments of 24° and 35°C, respectively; Supplementary Material Figure 4) (Tyler, 2011). Shielding the sensor (Dollberg et al., 1994), increasing the attachment thickness (Deng and Liu, 2008), or increasing the number of attachment layers (Tyler, 2011) tended to increase the measured temperature. One study with pre-term infants inside incubators had a pertinent risk of bias in that the measured T_{skin} itself influenced the environmental conditions within the incubator (Dollberg et al., 1994). In that study, higher measured temperature caused lower environmental temperatures and so this risk of bias likely mitigated the magnitude of the observed difference: indeed, the net effect was a 0.8°C greater skin-to-environment temperature gradient when the shielding attachment was used. Elsewhere, aluminium tape typically resulted in higher measured temperatures than other common surgical tapes (>1°C for pooled mean differences), although interestingly, these higher $T_{surface}$ were typically closer to the expected $T_{surface}$ of the aluminium plate (Psikuta et al., 2014) (see also section Thermal Equilibrium of the Surface Sensor with the Underlying Temperature).

Influence of the Pressure Applied by Surface Sensors

Data from four studies were identified (Yakovlev and Utekhin, 1966; Guadagni et al., 1972; Jirak et al., 1975; Mahanty and Roemer, 1979b) (n = 4 subsets; 2 subsets displayed, giving 43 comparisons; **Figure 6** and Supplementary Material Figure 5). Absolute mean differences ranged from 0 to 1.33°C for the pressure comparisons available (**Figure 6**). The measured temperature tended to increase with increasing pressure over the range of 2–681 mmHg (Jirak et al., 1975; Mahanty and Roemer, 1979b). Halving or doubling the applied pressure resulted in
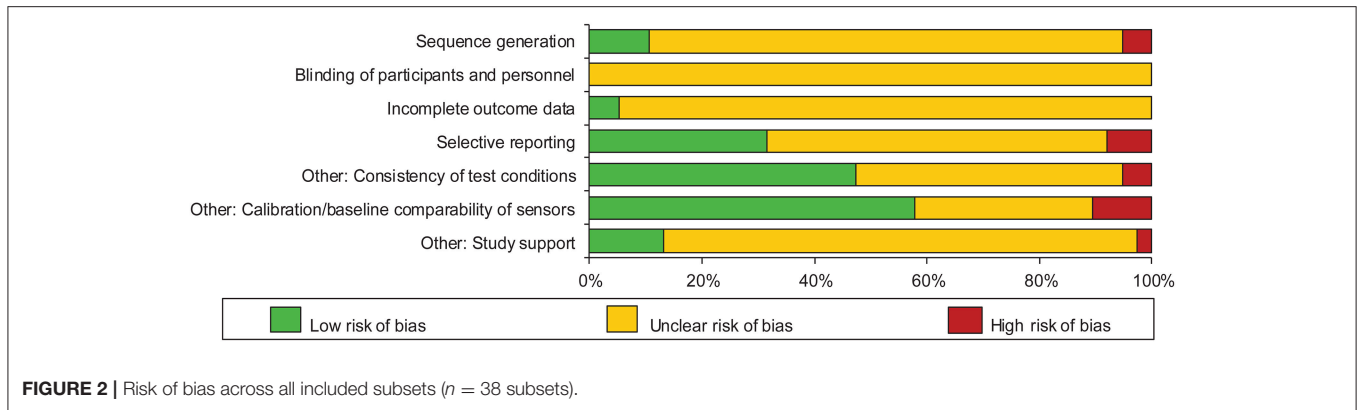
FIGURE 2 | Risk of bias across all included subsets ($n = 38$ subsets).
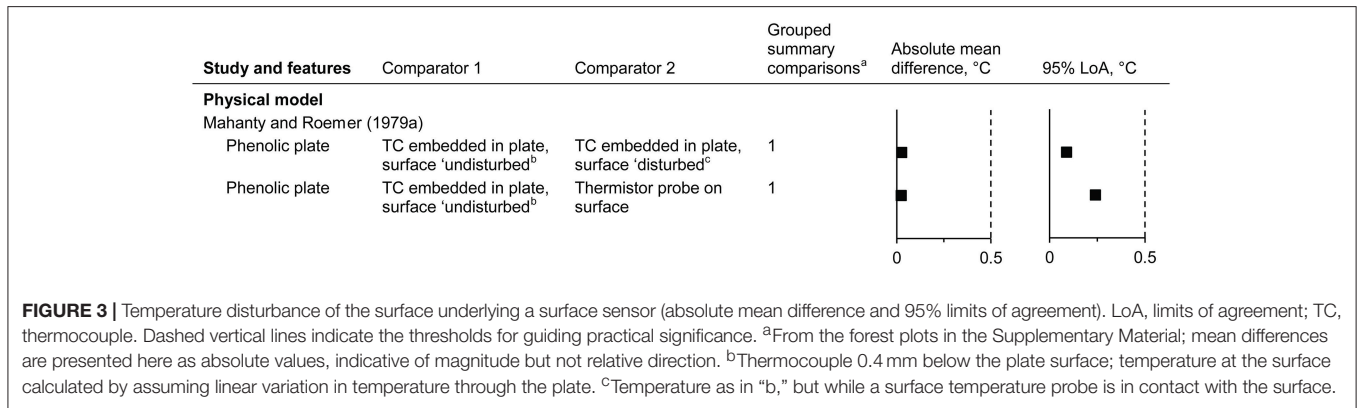

FIGURE 3 | Temperature disturbance of the surface underlying a surface sensor (absolute mean difference and 95% limits of agreement). LoA, limits of agreement; TC, thermocouple. Dashed vertical lines indicate the thresholds for guiding practical significance. [a]From the forest plots in the Supplementary Material; mean differences are presented here as absolute values, indicative of magnitude but not relative direction. [b]Thermocouple 0.4 mm below the plate surface; temperature at the surface calculated by assuming linear variation in temperature through the plate. [c]Temperature as in "b," but while a surface temperature probe is in contact with the surface.
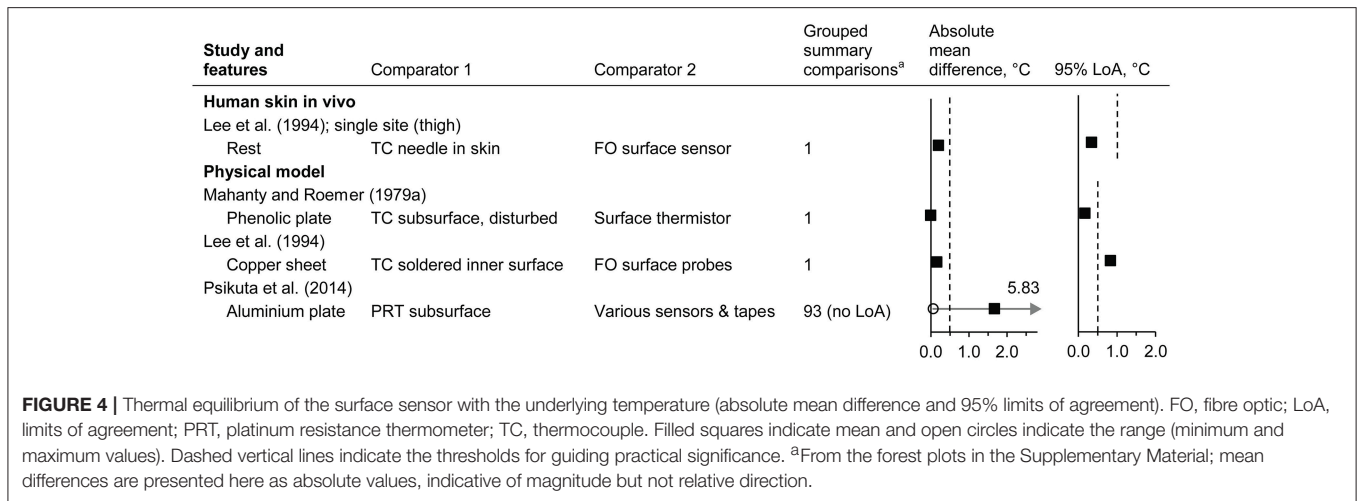

FIGURE 4 | Thermal equilibrium of the surface sensor with the underlying temperature (absolute mean difference and 95% limits of agreement). FO, fibre optic; LoA, limits of agreement; PRT, platinum resistance thermometer; TC, thermocouple. Filled squares indicate mean and open circles indicate the range (minimum and maximum values). Dashed vertical lines indicate the thresholds for guiding practical significance. [a]From the forest plots in the Supplementary Material; mean differences are presented here as absolute values, indicative of magnitude but not relative direction.

mean differences consistently within $\pm 0.5°C$ (note that the common comparison pressures were used as presented in the original articles) (Jirak et al., 1975; Mahanty and Roemer, 1979b). For example, for a circular sensor at the forehead the mean differences, versus 136 mmHg, were $-0.31°C$ at 68 mmHg and $+0.36°C$ at 272 mmHg (Jirak et al., 1975; Supplementary Material Figure 5). From the data of one subset (Jirak et al., 1975), the LoA were typically greater when the higher pressures were included (409, 545, and 681 mmHg), although all within $\pm 1.0°C$.

One of the subsets not able to be displayed in the plots (Guadagni et al., 1972) was simply summarised in the original article as a statement that the effect of pressure on steady state skin temperature (human participants; sites not reported) was $<0.01°C$ at pressures between 31 and 52 mmHg, and that at pressures beyond this range, the steady state temperature becomes pressure-dependent and increases with increasing pressure. For the other subset not able to be displayed (Yakovlev and Utekhin, 1966), differences of $0.1–0.7°C$ were reported for
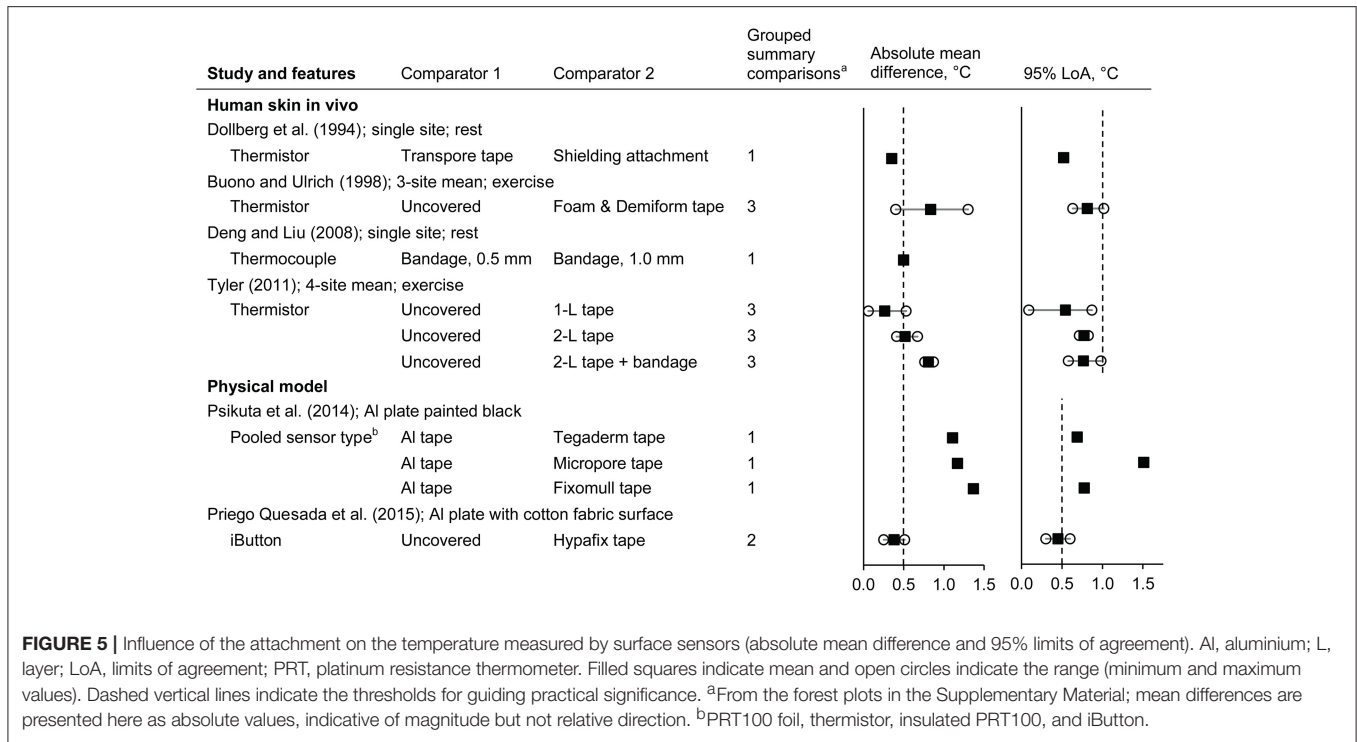
FIGURE 5 | Influence of the attachment on the temperature measured by surface sensors (absolute mean difference and 95% limits of agreement). Al, aluminium; L, layer; LoA, limits of agreement; PRT, platinum resistance thermometer. Filled squares indicate mean and open circles indicate the range (minimum and maximum values). Dashed vertical lines indicate the thresholds for guiding practical significance. [a]From the forest plots in the Supplementary Material; mean differences are presented here as absolute values, indicative of magnitude but not relative direction. [b]PRT100 foil, thermistor, insulated PRT100, and iButton.
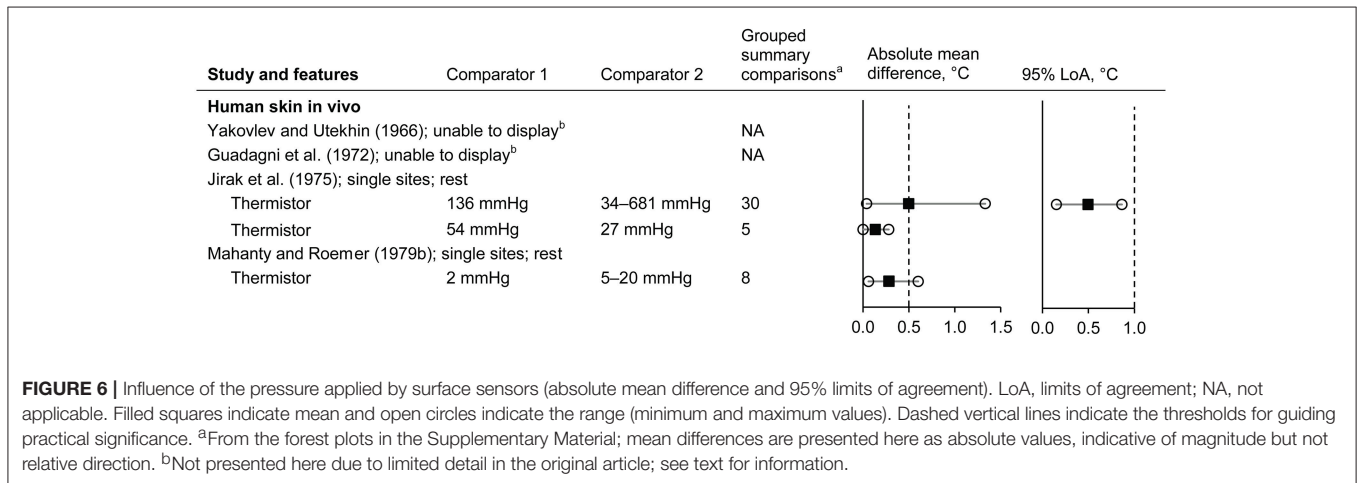

FIGURE 6 | Influence of the pressure applied by surface sensors (absolute mean difference and 95% limits of agreement). LoA, limits of agreement; NA, not applicable. Filled squares indicate mean and open circles indicate the range (minimum and maximum values). Dashed vertical lines indicate the thresholds for guiding practical significance. [a]From the forest plots in the Supplementary Material; mean differences are presented here as absolute values, indicative of magnitude but not relative direction. [b]Not presented here due to limited detail in the original article; see text for information.

the seven sensor types tested on human skin (sites not reported) whereby the pressure was generated by 5 and 70 g weights (sensor surface areas were not reported, therefore the pressures are unknown). Notwithstanding, the authors did state, without the supporting data, that pressure effects can be disregarded when the maximum pressure of the sensor on the skin does not exceed 15–37 mmHg.

Influence of the Environmental Conditions on Surface Sensors

Data from six studies were identified (Yakovlev and Utekhin, 1966; Buono and Ulrich, 1998; van Marken Lichtenbelt et al., 2006; Harper Smith et al., 2010; Tyler, 2011; Psikuta et al., 2014)

($n = 6$ subsets; five subsets displayed, giving 38 comparisons; **Figure 7** and Supplementary Material Figure 6).

For the subsets in which human participants were used (Buono and Ulrich, 1998; van Marken Lichtenbelt et al., 2006; Harper Smith et al., 2010; Tyler, 2011), potential influences of the environment can be examined indirectly via the relative differences for the same two comparators under different environments—here, the different environments cannot be directly compared because the T_{skin} will also change. (Because these indirect comparisons may be confounded by other variables such as the differences in absolute T_{skin}, they are interpreted cognizant of this possible but unclear risk of bias.) Irrespective of the underlying cause, the mean difference between the same two
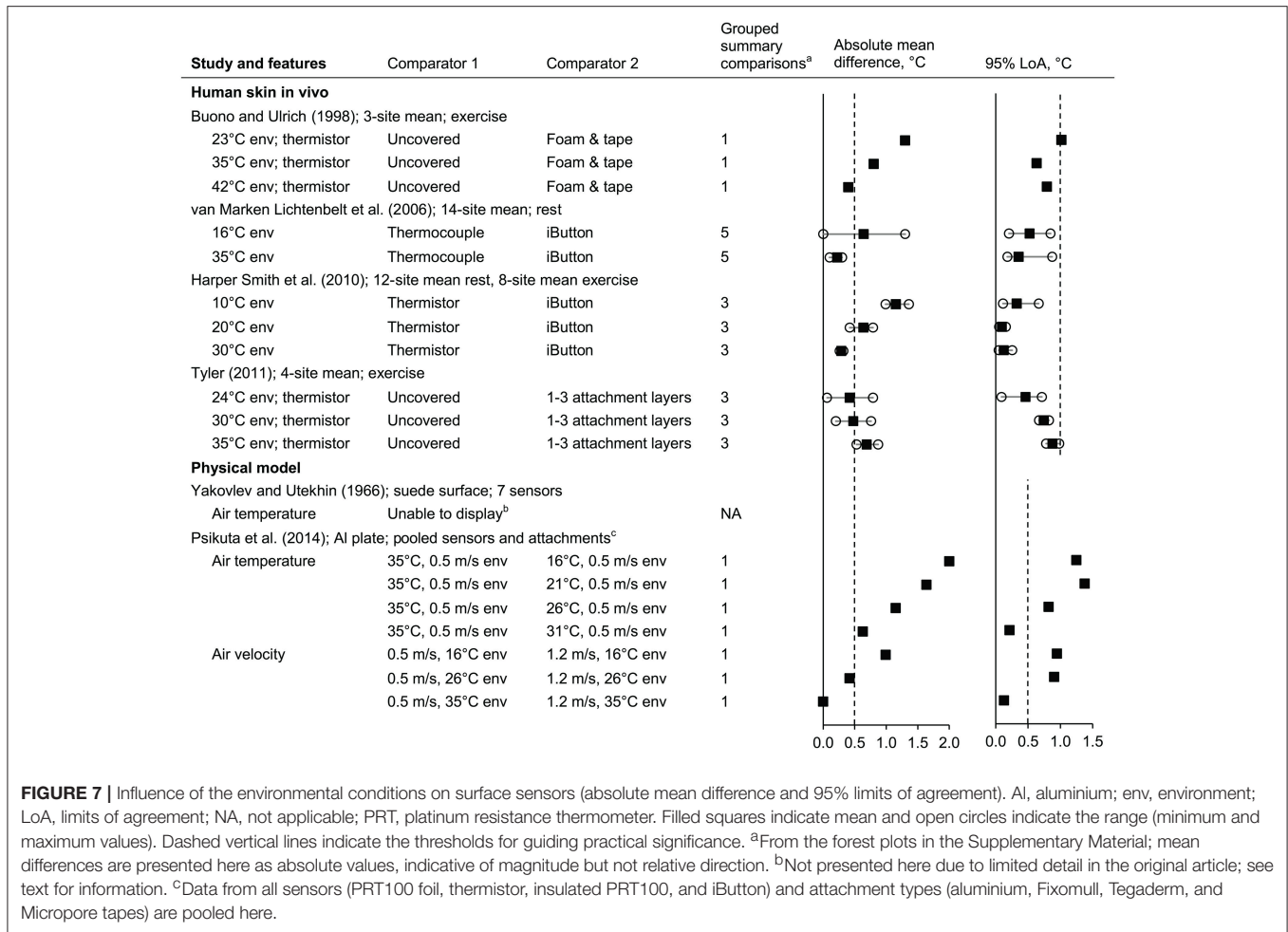
FIGURE 7 | Influence of the environmental conditions on surface sensors (absolute mean difference and 95% limits of agreement). Al, aluminium; env, environment; LoA, limits of agreement; NA, not applicable; PRT, platinum resistance thermometer. Filled squares indicate mean and open circles indicate the range (minimum and maximum values). Dashed vertical lines indicate the thresholds for guiding practical significance. [a]From the forest plots in the Supplementary Material; mean differences are presented here as absolute values, indicative of magnitude but not relative direction. [b]Not presented here due to limited detail in the original article; see text for information. [c]Data from all sensors (PRT100 foil, thermistor, insulated PRT100, and iButton) and attachment types (aluminium, Fixomull, Tegaderm, and Micropore tapes) are pooled here.

comparators were often different when the comparison was made under different environmental temperatures in three subsets (Buono and Ulrich, 1998; van Marken Lichtenbelt et al., 2006; Harper Smith et al., 2010): for example, relative mean differences of 1.3 and 0.4°C for the same attachment comparison under 23 and 42°C environmental temperatures (Buono and Ulrich, 1998). In other words, the magnitude of mean difference for a given comparison of sensor setups showed some dependence on the environmental temperature. For the influence of air movement during rest, the pattern of relative differences at 10, 20, and 30°C environmental temperatures was similar under both 0.2 and 2.3 m/s air velocities (e.g., relative mean differences of 0.70 and 0.79°C for the two velocities, respectively, in a 20°C environment; Supplementary Material Figure 6) (Harper Smith et al., 2010).

For one subset in which a model was used (Psikuta et al., 2014), data are presented in **Figure 7** as pooled estimates for main effects of the environmental variables. In this subset (Psikuta et al., 2014), compared to $T_{surface}$ measurements in the 35°C environment, the cooler environments resulted in lower, and more variable, measured values despite the plate being maintained at ~36.5°C; the magnitudes of mean differences were often of practical relevance (>0.5°C). Comparing wind velocity within the same environmental temperature indicates that the influence of an increase in wind velocity (from 0.5 to

1.2 m/s) was negligible when the environmental temperature (35°C) was close to the $T_{surface}$ (36.5°C) but became increasingly relevant as the environmental temperature diverged further from the plate temperature (relative differences of −0.4 and −1.0°C for the point estimates in environments of 26 and 16°C, respectively). Data from the other subset using a physical model are not presented due to uncertainties in interpretation of the environment and the reference $T_{surface}$ but the authors did state effects of the ambient air temperature being 0.2–6.0°C for seven different sensor types (Yakovlev and Utekhin, 1966).

Influence of the Type of Surface Sensor

Data from 14 studies were identified (Yakovlev and Utekhin, 1965, 1966; Jirak et al., 1975; Flesch et al., 1976; Mahanty and Roemer, 1979a,b; Krause, 1993; van Marken Lichtenbelt et al., 2006; Harper Smith et al., 2010; Youhui et al., 2010; James et al., 2014; Psikuta et al., 2014; Bach et al., 2015a; McFarlin et al., 2015) ($n = 17$ subsets; 14 subsets displayed, giving 87 comparisons; **Figure 8** and Supplementary Material Figure 7). The risk of bias for calibration of these sensors used was considered high for four subsets [*in vivo* (McFarlin et al., 2015) and models (Flesch et al., 1976; Krause, 1993), (James et al., 2014) model uncorrected], unclear for five subsets [*in vivo* (Yakovlev and Utekhin, 1965, 1966; Mahanty and Roemer, 1979a; Youhui et al., 2010), and
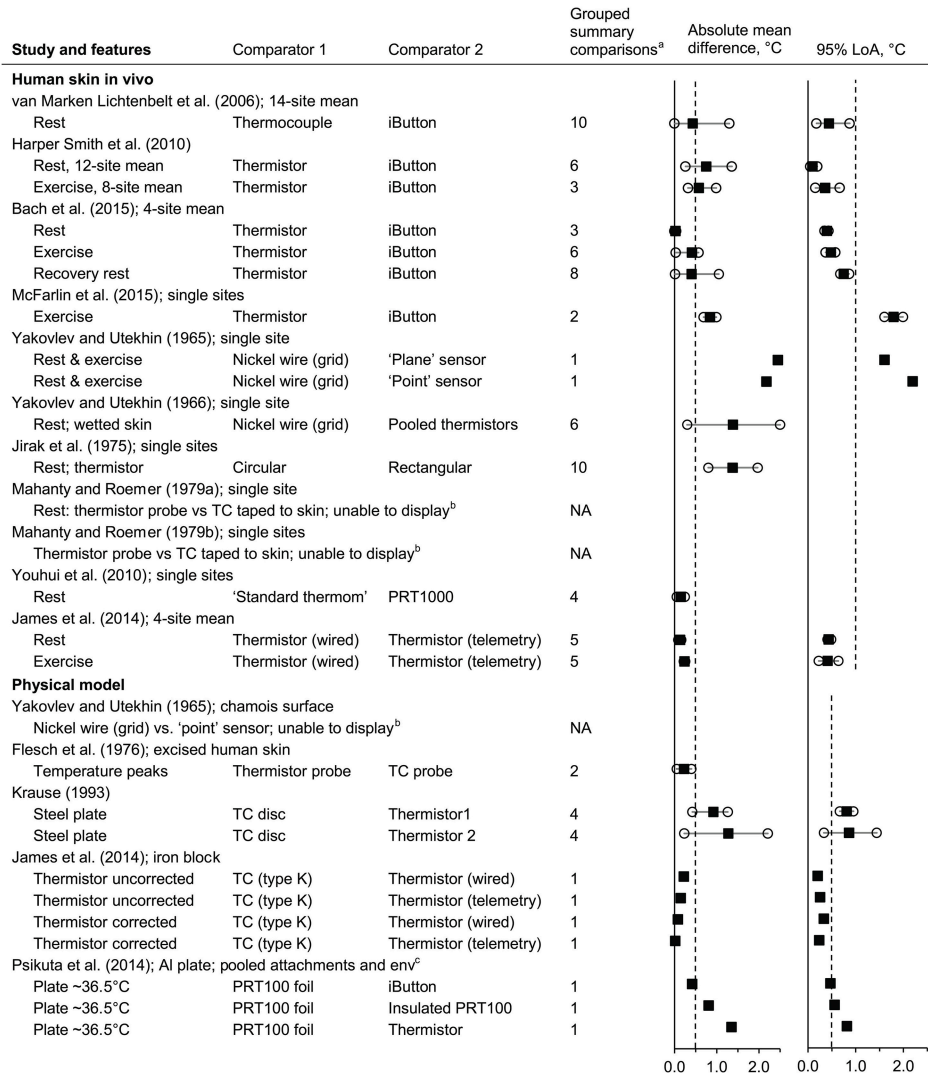
Study and features	Comparator 1	Comparator 2	Grouped summary comparisons[a]	Absolute mean difference, °C	95% LoA, °C
Human skin in vivo					
van Marken Lichtenbelt et al. (2006); 14-site mean					
Rest	Thermocouple	iButton	10		
Harper Smith et al. (2010)					
Rest, 12-site mean	Thermistor	iButton	6		
Exercise, 8-site mean	Thermistor	iButton	3		
Bach et al. (2015); 4-site mean					
Rest	Thermistor	iButton	3		
Exercise	Thermistor	iButton	6		
Recovery rest	Thermistor	iButton	8		
McFarlin et al. (2015); single sites					
Exercise	Thermistor	iButton	2		
Yakovlev and Utekhin (1965); single site					
Rest & exercise	Nickel wire (grid)	'Plane' sensor	1		
Rest & exercise	Nickel wire (grid)	'Point' sensor	1		
Yakovlev and Utekhin (1966); single site					
Rest; wetted skin	Nickel wire (grid)	Pooled thermistors	6		
Jirak et al. (1975); single sites					
Rest; thermistor	Circular	Rectangular	10		
Mahanty and Roemer (1979a); single site					
Rest: thermistor probe vs TC taped to skin; unable to display[b]			NA		
Mahanty and Roemer (1979b); single sites					
Thermistor probe vs TC taped to skin; unable to display[b]			NA		
Youhui et al. (2010); single sites					
Rest	'Standard thermom'	PRT1000	4		
James et al. (2014); 4-site mean					
Rest	Thermistor (wired)	Thermistor (telemetry)	5		
Exercise	Thermistor (wired)	Thermistor (telemetry)	5		
Physical model					
Yakovlev and Utekhin (1965); chamois surface					
Nickel wire (grid) vs. 'point' sensor; unable to display[b]			NA		
Flesch et al. (1976); excised human skin					
Temperature peaks	Thermistor probe	TC probe	2		
Krause (1993)					
Steel plate	TC disc	Thermistor1	4		
Steel plate	TC disc	Thermistor 2	4		
James et al. (2014); iron block					
Thermistor uncorrected	TC (type K)	Thermistor (wired)	1		
Thermistor uncorrected	TC (type K)	Thermistor (telemetry)	1		
Thermistor corrected	TC (type K)	Thermistor (wired)	1		
Thermistor corrected	TC (type K)	Thermistor (telemetry)	1		
Psikuta et al. (2014); Al plate; pooled attachments and env[c]					
Plate ~36.5°C	PRT100 foil	iButton	1		
Plate ~36.5°C	PRT100 foil	Insulated PRT100	1		
Plate ~36.5°C	PRT100 foil	Thermistor	1		

0.0   1.0   2.0   0.0   1.0   2.0

FIGURE 8 | Influence of the type of surface sensor (absolute mean difference and 95% limits of agreement). Al, aluminium; env, environment; LoA, limits of agreement; NA, not applicable; PRT, platinum resistance thermometer; TC, thermocouple; thermom, thermometer. Filled squares indicate mean and open circles indicate the range (minimum and maximum values). Dashed vertical lines indicate the thresholds for guiding practical significance. [a]From the forest plots in the Supplementary Material; mean differences are presented here as absolute values, indicative of magnitude but not relative direction. [b]Not presented here due to the limited information reported in original article; see text for information. [c]Data from all attachment types (aluminium, Fixomull, Tegaderm, and Micropore tapes) and environments (15–35°C, 0.5 m/s) are pooled here.

model (Yakovlev and Utekhin, 1965)], and low for eight subsets [*in vivo* (Jirak et al., 1975; Mahanty and Roemer, 1979b; van Marken Lichtenbelt et al., 2006; Harper Smith et al., 2010; James et al., 2014; Bach et al., 2015a) and models (Psikuta et al., 2014; James et al., 2014) model corrected].

Absolute mean differences for some subsets were almost consistently within (Youhui et al., 2010; James et al., 2014) or beyond 0.5°C (Yakovlev and Utekhin, 1965, 1966; Jirak et al., 1975; Krause, 1993; McFarlin et al., 2015). For four *in vivo* studies, the LoA were almost consistently within ±1.0°C (van Marken Lichtenbelt et al., 2006; Harper Smith et al., 2010; James et al., 2014; Bach et al., 2015a) and in two studies consistently greater than ±1.0°C (Yakovlev and Utekhin, 1965; McFarlin et al., 2015).

Data from three subsets were not able to be presented in the plots due to insufficient information [human participants at rest (Mahanty and Roemer, 1979a,b) and a model (Yakovlev and Utekhin, 1965)]. According to the text of the original articles, a thermocouple taped to the skin and a thermistor probe "agreed within ±0.15°C (maximum difference)" on one occasion (Mahanty and Roemer, 1979a) and had a mean difference of 0.03°C on the other occasion (Mahanty and Roemer, 1979b). For the physical model (Yakovlev and Utekhin, 1965), a nickel wire sensor and a "point" sensor (diameter 0.3–1 mm) were compared following the wetting of the suede surface of a model: the temperature from the nickel wire sensor decreased by approximately 4°C over the subsequent 2 min whereas the

temperature from the "point" sensor decreased by <1°C over the same period.

DISCUSSION

In comparison measurements from various contact T_{skin} sensor setups—using human participants or relevant physical models—mean differences exceeding 0.5°C were observed for some comparisons of attachment, applied pressure, and sensor type. Additionally, there was indication that the environmental conditions can influence the measured temperature and that the surface sensor is not always sufficiently at thermal equilibrium with the underlying surface. Thus, the sensor type used and how it is used can meaningfully influence the measured value. For users of contact T_{skin} sensors, emphasis should be placed on consistency of sensor setup parameters (at least within a study) and the limitations of knowing an absolute T_{skin} must be recognised, particularly for comparability across studies. The survey of contact T_{skin} measurements here illustrates that, for the range of sensor setups used in different studies involving sport, exercise, and other PA, important information about features like calibration and the attachment method are often unreported. For the continuing improvement of research and clinical measurements, and with growing commercial and consumer interest in wearable technology and personal body monitoring, this review demonstrates that even routine measurements are not necessarily as simple as they may otherwise seem.

Measurement Comparisons
Strengths and Limitations
This is the first systematic review to comprehensively collate data from comparisons of measurements made using contact T_{skin} sensors. A strength of this work is the breadth of information included and the ability to examine specific outcomes of interest in detail.

While comparisons of outcome data were typically possible, meta-analyses were not performed due to the appreciable variety of test conditions, sensor setups, and types of temperature data used (e.g., single site versus mean T_{skin}; multiple measurements within a period versus period grand mean), and the resultant lack of a common reference condition. Further, the lack of a universal gold-standard method for measuring T_{skin} limited the ability to judge one comparator T_{skin} sensor setup as being more accurate than the other. For this reason, emphasis was typically placed on the magnitude of effects rather than the direction of the effect.

The often limited reporting of methodological information contributed to the number of "unclear" judgements within the risk of bias assessments and made extraction of study information or data challenging at times. To guide both study design and reporting of study information in future studies comparing temperature measurements, we recommend authors to prospectively think about how their work will be judged during a subsequent risk of bias assessment. Details to consider include—but are not limited to—sequence generation, any blinding, completeness of outcome data (e.g., detachment of sensors, sensor malfunctions), selective reporting (data that was intended to be reported but was not), consistency of test conditions, sensor calibration details, and study support.
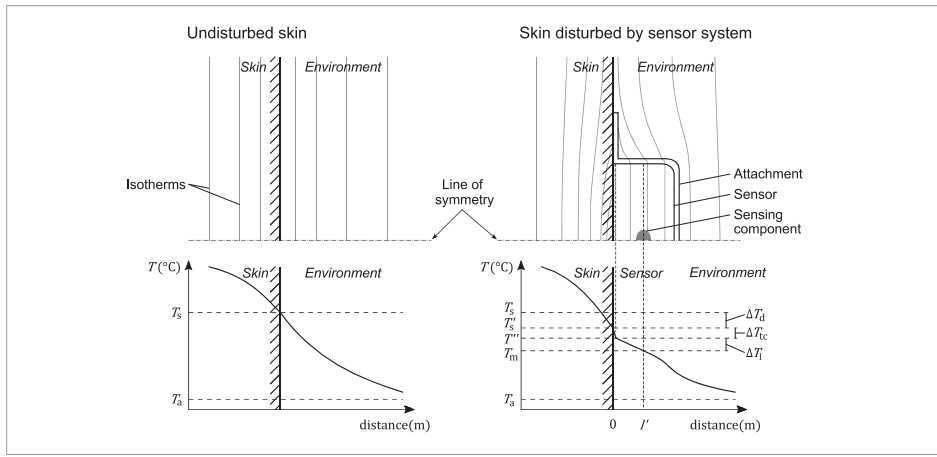
We used threshold values to indicate practical significance throughout this review (human skin *in vivo*, ±0.5°C for mean difference and ±1.0°C for 95% LoA; physical models, ±0.5°C for mean difference and 95% LoA). This was a necessary simplification for feasibility and is intended as a general guide only. It is likely that, in practice, these thresholds will need to be adjusted according to specific measurement contexts and objectives. For example, multiple spot measurements from contact T_{skin} sensors are often used to calculate a mean T_{skin}, ostensibly representative of the body as a whole. Temperature differences due to the sensor setup itself may in some cases be within differences due to other experimental decisions such as site selection, replication of sensor placement, or the weightings used (Livingstone et al., 1987; Choi et al., 1997).

It is acknowledged that additional descriptive statistics are also useful for making inferences in method comparisons (e.g., typical error of the estimate, regression and correlation coefficients). Finally, not considered here was any influence of the signal processor/data logger setup (Jutte et al., 2005).

Validity of the Surface Measurement
While the skin surface is an accessible site for measurement, it presents difficulties for accurate quantification of temperature in that the surface is an interface between distinct mediums, with each medium having its own thermal properties and temperature gradients. Accordingly, the sensor system is in partial contact with medium of interest (i.e., the skin) and partial contact with the adjacent environment (e.g., microclimate air, liquids, clothing). Accurate knowledge of T_{skin} requires the sensor temperature to be in equilibrium with the corresponding skin temperature and for the sensor setup itself to have negligible additional effect on heat exchanges between the underlying skin and its adjacent environment(s) (Bedford and Warner, 1934; Hardy, 1934b; Childs, 2001; Taylor et al., 2014). Demonstrable differences among various T_{skin} sensor setups may reduce, at least in part, to how well each respective sensor setup achieves this balance between thermal equilibrium and temperature modification. Selected sources of error that may be expected during T_{skin} measurements are given in **Figure 9**, along with suggestions for how such an error may be minimised (Michalski et al., 2001; Nicholas and White, 2001). Considering sources of error and error minimisation can assist T_{skin} sensor design or assist sensor selection for those measuring T_{skin}. Notwithstanding, ideal dimensions or thermal properties of sensors and their attachments cannot govern sensor design or setup selection alone as these devices must also be robust and suitable enough for practical use (Youhui et al., 2010; Webb et al., 2013).

The data identified within this review were limited and incompletely address the question of validity of common T_{skin} sensors and so future work is warranted. In particular, estimates of uncertainty during T_{skin} measurements and the magnitude of any effects of the sensor and attachment disturbing the underlying T_{skin} are unclear. From data of the only study identified, there was no demonstrable effect of thermal disturbance of the site underlying a surface sensor (Mahanty and Roemer, 1979a). However, it is unlikely this finding is

FIGURE 9 | Measurement of skin temperature using a contact sensor attached to the skin surface is associated with a number of potential sources of error, some of which are illustrated and described here (Michalski et al., 2001; Nicholas and White, 2001). The schematic represents, in the case of undisturbed skin (left) and disturbed skin (right), cross sections with hypothetical isotherms (top) and the corresponding temperature profiles at the line of symmetry (bottom).

generalizable to conventional T_{skin} sensor types and across environmental conditions as the surface probe used was designed specifically to minimise any such modification of $T_{surface}$ and the environmental temperature used ($\sim$22°C) represented a moderate difference from the surface temperature (29–33°C). Notwithstanding, this finding does illustrate the promise of sensor design based on error minimisation (**Figure 9**). Key characteristics of this sensor included using a small sensing component (thermistor) housed in a stainless steel disc (19 mm diameter, 0.13 mm thick), which was partially painted matte

black (high emissivity, similar to skin) and partially covered with reflective silver coating to compensate for addition heat losses due in part to the surface area of the sensor system being greater than that of the skin itself (Mahanty and Roemer, 1979a).

Numerical modelling has indicated that insulated T_{skin} sensors can cause the temperature of the underlying skin to be higher than it would otherwise be because insulating the sensor from the adjacent environment also insulates the underlying skin (Boetcher et al., 2009). On the contrary, sensors with high thermal conductivity and no insulating layer can lower the

underlying T_{skin} by inducing a fin effect whereby the total surface area of the sensor exposed to the environment is greater than that of the skin it covers and, in this way, heat conducted from the skin to the sensor can be lost by convection to the adjacent environment at a rate that exceeds that of the undisturbed skin itself (Guadagni et al., 1972; Boetcher et al., 2009). Heat transfer from the sensor to the environment by radiation may be greater or lesser than that of the undisturbed skin because, despite the greater surface area available for radiating heat in the case of the sensor, the emissivity of the skin is itself very high (∼0.98; Steketee, 1973).

The skin has a boundary layer of air that mitigates the temperature difference directly adjacent to its surface. This difference under warm-hot conditions (relatively still air, ambient temperature of 28–29°C) may be <0.5°C for air within 1 mm, and <1°C for air up to 6 mm from the skin surface (McGlone and Bazett, 1927). In cooler conditions or with forced air convection (e.g., wind or body movements), this gradient increases and the temperature difference can become appreciable within the height of commonly used T_{skin} sensors and their attachments. For surface T_{skin} sensor equilibrium with the underlying temperature, a range of magnitudes was observed from <0.5°C (Mahanty and Roemer, 1979a; Lee et al., 1994; Psikuta et al., 2014) up to 6°C (Psikuta et al., 2014). However, details about the ambient environment were not clear from the two of the studies included here (Mahanty and Roemer, 1979a; Lee et al., 1994).

Psikuta et al. (2014) observed that ambient temperature, and thus the surface-environment gradient, can influence the estimated error: while few sensor-attachment combinations were within 1°C of the plate temperature in the 16°C environment, all were within this range in the 35°C environment. The temperature of the aluminium plate was reported as 36.5±0.03°C in this experiment and, therefore, it is intuitive that errors tend to reduce as the ambient temperature approaches the surface temperature.

Although numerical modelling and non-contact methods can also add insight, it was not feasible to also cover these in detail here and contact versus non-contact methods have been discussed elsewhere (Bach et al., 2015b). As mentioned previously, numerical modelling has indicated that, at a simulated environmental temperature of 22°C and T_{skin} of ∼31.7°C, errors approaching 0.5°C can be expected with insulated contact sensors (Boetcher et al., 2009). Numerically modelling performed by Deng and Liu (2008) indicated that, for the attachment material, decreasing the thickness and increasing the thermal conductivity decreased the measurement error. These results are consistent with experimental observation of aluminium tape tending to better reflect the underlying plate temperature, particularly at low wind velocities (0.5 m/s) (Psikuta et al., 2014). For improving absolute temperature measurement, more work is required to delineate measurement errors in the context of human skin, especially during sweating.

Comparisons of Setup Variables

Sensor setup variables can influence the measured values from T_{skin} sensors. Examples of both small and practically significant effects (less than or greater than 0.5°C absolute mean difference, respectively) can be found for the outcome subgroups of attachment type, sensor-applied pressure, environmental conditions, and sensor type. The 95% LoA were often, though not always, within ±1.0°C for human skin *in vivo* and ±0.5°C for physical models. An implication of these observations is that equivalent measurements of T_{skin} cannot be assumed when different components of the sensor setup or different environments are used. Thus, these setup variables (e.g., sensor type, attachment, environmental conditions) should be considered prior to measurement and clearly reported to assist subsequent interpretation of T_{skin} measurements. Further work could identify or experimentally justify particular setups or conditions in which measurement biases are mitigated (**Figure 9**).

Relative temperature differences between sensor types can vary with the environmental temperature, activity, and time (van Marken Lichtenbelt et al., 2006; Harper Smith et al., 2010; Bach et al., 2015a), although this is not always the case (James et al., 2014). In other words, two different setups may not simply be a consistent offset from each other, but rather vary in magnitude with other parameters. Physical properties of the sensors and their attachments likely contribute to the consistency, or inconsistency, of relative differences, particularly during transients.

Some patterns in temperature differences during changes in environmental temperatures could reflect an influence of thermal inertia (Supplementary Material Figure 7A): iButtons appeared to either have a reduced response time to a true change in T_{skin} or, alternatively, they were less influenced by the change in environment and better reflected the true change in T_{skin} (van Marken Lichtenbelt et al., 2006; Bach et al., 2015a). The practical consequences due to the thermal inertia of the sensor setup will likely depend on the time course of the expected change in T_{skin}. The response time of the sensor on the skin surface should be a key consideration when relatively fast changes in T_{skin} are expected (Arunachalam et al., 2008), but may be of lower importance compared to other considerations when a relatively stable T_{skin} is expected. In research settings it is common—and good practice—for participant instrumentation to occur well before experimental recordings are made, which should mitigate any transient effects of initial sensor temperature before application and the duration of a sensor to reach a thermal steady state once applied (Guadagni et al., 1972).

Surface pressure is known to influence skin perfusion, however, studies here indicated that the measured T_{skin} temperature almost invariably increased with increasing pressure whereas skin perfusion tends to decrease, even with a little as 5 mmHg (Holloway et al., 1976). Physical effects of the surface sensor pressing into the skin were investigated further by Mahanty and Roemer (1979b) using numerical modelling and the calculated effects agreed closely with their experimental results, suggesting these physical factors associated with depression are a suitable explanation. While increasing the sensor-applied pressure may lead to temperature errors, the thermal contact must also be sufficient and, therefore, there is likely a trade-off between the two. Unstable temperature readings were

attributed to insufficient thermal contact at pressures lower than 31 mmHg in one study (Guadagni et al., 1972), but such instability issues were not apparent in another study investigating pressures of 2–20 mmHg (Mahanty and Roemer, 1979b). In the latter study, the authors suggested that 2 mmHg was sufficient for measuring T_{skin} and used numerical modelling to support this observation (Mahanty and Roemer, 1979b).

Controlling the applied pressure is a challenge for measuring T_{skin} *in vivo*. While sensor-applied pressure variations during normal use are not clear, it is notable that mean differences in T_{skin} due to halving or doubling a given applied pressure were <0.5°C. This finding was consistent across the wide range of pressures investigated (2–681 mmHg). Increasing the sensor surface area can be considered with respect to limiting the applied pressure and associated depression into the skin, but any such increases in area may have consequences for modifying evaporative and non-evaporative heat exchanges between the skin and the environment.

A challenge for measurement comparisons on human skin is that T_{skin} varies by time and location (Pennes, 1948; Kaufman and Pittman, 1966; Webb, 1992). Possible comparisons on human skin include using identical measurement sites at different times, or simultaneous measurements at different sites. While some authors compare results from single sites, others use multiple measurement sites and compare the resultant mean T_{skin}. With respect to temporal changes in T_{skin} of human participants, Guadagni et al. (1972) observed that after 2.5–3 h rest in a room temperature of 24.5°C, changes in T_{skin} at a given site were <0.01°C within 2 or 3 min periods. A different approach was used by Mahanty and Roemer (1979b) whereby throughout a measurement period involving different experimental sensor-applied pressures, adjacent temperature sensors remained in place to allow minor temperature changes not due to experimental sensor pressure to be compensated for. Irrespective of the approach, care needs to be taken to minimise risks of systematic temperature bias.

With the expanding interest in wearable technology for day-to-day applications, there should be an expectation of acceptable validity for any measurements being taken (Sperlich and Holmberg, 2017). A direct outcome here is that sensor location and the effects of the carrier devices or attachments will need to be taken into account for such applications.

Survey of Use
Strengths and Limitations
Here we have collected information about how contact T_{skin} sensors are currently (2011–2016) used and reported in published studies involving sport, exercise and other PA. Thus, this work provides context and a basis for integrating the findings delineated above in more challenging conditions than rest. These findings may, however, only provide limited information on these aspects in other research or clinical conditions.

General Implications
In human studies, details about T_{skin} sensor setup details were often poorly reported and, from those reporting setup information, it was evident that the setups often varied in terms of the sensors, attachments, and locations used. Key setup variables need to be considered further and consistently reported. For example, the sensor attachment method was clear in only 42% of the articles sampled yet attachment can bias the measured values (Buono and Ulrich, 1998; Psikuta et al., 2014). Similarly, clear information about sensor calibration was present in only 6% these studies.

Absolute T_{skin} values are used often in research with 97% of the studies sampled here reporting some form of absolute T_{skin} and 17% reporting change scores (some reported both, but not necessarily for all the data). While common inferential statistical approaches end up treating absolute values as relative differences (e.g., within-subject differences for repeated measures, difference between group means for independent samples), the point remains that absolute values are of interest to researchers and other users of T_{skin} information. Thus, clearer reporting of setup variables may also improve awareness about the external validity of published absolute T_{skin} data, such as for inter-study comparisons or applied use of T_{skin} (e.g., temperature thresholds during heat stress).

Sensors used in the comparisons of measured temperatures (objective 1) only partially represented the sensors currently used in human studies involving PA. The outcome of some comparisons may be more generally applicable, such as effects of the sensor pressure, whereas the outcomes of other comparisons may be restricted to specific setups or environments, such as sensor type or a particular skin-environment temperature gradient. A reason contributing to specificity of comparison outcomes is that different sensor setups likely vary in variables not taken into account here, such as sensor and attachment composition (and resultant thermal properties), dimensions, and surface contact. Future work will benefit from the identification of setup variables suitable for widespread use.

Summary
This work serves to establish a base from which users of contact T_{skin} measurements can make better-informed decisions about setup of the sensor-attachment system with relevance to their particular measurement objectives and the expected measurement conditions. Some potentially relevant components within the measurement system are summarised in **Table 2**.

For those performing studies comparing different T_{skin} sensors or sensor setup variables (e.g., method agreement studies), key points of consideration include:

- There is currently no universally applicable gold-standard T_{skin} measurement, which likely contributes to the general lack of a common reference condition across studies. In lieu of a specific sensor setup available to be used consistently among studies, key features of the setups used need to be described. This information includes the sensor manufacturer and model, type (e.g., thermistor), some physical characteristics (e.g., insulated or uninsulated, physical dimensions), attachment, and calibration details.
- We recommend that authors consider criteria used in risk of bias assessments to help guide study design and reporting.

TABLE 2 | Summary of some relevant considerations for the measurement of skin temperature.

Component of system	Consideration for component
Skin	• Selection of sites and consistency of placement • Homogeneity of surface skin temperature • Temperature gradients within the skin itself • Lateral heat transfer from skin adjacent to the site being measured • Sweating • Sensor pressure (depression into the skin)
Skin-sensor interface	• Thermal contact resistance • Contact surface area
Sensor	• Physical dimensions (including surface area of the sensor versus that of the skin site covered) and thermal properties (including the resultant thermal resistance) • Obstruction of evaporative heat loss from skin • Calibration method and range
Attachment	• Physical dimensions and thermal properties (including thermal resistance and water vapour resistance) • For tapes covering the sensor, emissivity and wetting characteristics for distributing liquid sweat • Effectiveness of remaining in place during prolonged use or physical activity
Clothing or coverings and associated microclimate	• Present or not • Temperature gradients (microclimate temperature will typically be closer to skin temperature where clothing/coverings used than without) • Air velocity • Humidity
Wider environment for uncovered sites	• Temperature gradients (difference from skin temperature) including air temperature and radiant temperature • Air velocity • Humidity

This list is not complete and not all of the factors will be relevant in each case.

This information includes sequence generation, any blinding, completeness of outcome data (e.g., detachment of sensors, sensor malfunctions), selective reporting (data that was intended to be reported but was not), consistency of test conditions, sensor calibration details, and study support.

• There is scope for future work to better clarify the measurement validity of commonly used T_{skin} sensor setups, particularly under conditions in which the skin-to-environment gradient becomes large and when relatively fast changes in T_{skin} are expected. Practical effects of surface T_{skin} sensors on modification of heat and mass transfer from the skin surface require experimental investigation.

For those measuring T_{skin} for research, clinical, or applied purposes, points of consideration include:

• Accurate quantification of an absolute T_{skin} (e.g., 35.5°C) is difficult and caution needs to be applied when taking and interpreting such data, particularly when comparing T_{skin} across studies when the method of measurement varies.

• Equivalent measurements of T_{skin} cannot be assumed when different components of the sensor setup or different environments are used. Thus, for sensor setup, emphasis should be placed on consistency of sensor setup parameters, at least within a study. This consistency includes the sensor used, attachment, and (as much as practicable) the sensor-applied pressure.

• Special care needs to be taken when the temperature gradients within the measurement system are high (e.g., low environment temperature) because errors can become more influential.

• As above, reporting sensor setup variables, calibration information, and clear site details better facilitates the interpretation and external use of T_{skin} data.

• The information in **Figure 9** may be used to assist sensor selection based on principles of minimising errors.

• In the design of future T_{skin} sensor systems, consideration should be given to retaining the practicality of simple affixable sensor systems while mitigating any temperature effects associated with the modification of skin coverage.

CONCLUSION

Contact T_{skin} sensor setups and conditions used vary considerably and reporting of this information is often incomplete. The range of measurement comparisons examined here indicated that (1) the basic validity of commonly used surface T_{skin} sensors for accurately measuring T_{skin} remains unclear, and (2) the sensor type used and how it is used can meaningfully influence the measured value. Key setup variables need to be appropriately considered and consistently reported.

AUTHOR CONTRIBUTIONS

All authors made contributions to the conception or design of this work and/or acquisition, synthesis, and interpretation of the data/information. BM drafted the manuscript with critical review from SA, CS, and RR. All authors approve the final version to be published and agree to be accountable for all aspects of the work.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphys. 2018.00029/full#supplementary-material

REFERENCES

Adams, J. D., Ganio, M. S., Burchfield, J. M., Matthews, A. C., Werner, R. N., Chokbengboun, A. J., et al. (2015). Effects of obesity on body temperature in otherwise-healthy females when controlling hydration and heat production during exercise in the heat. *Eur. J. Appl. Physiol.* 115, 167–176. doi: 10.1007/s00421-014-3002-y

Adams, J. D., McDermott, B. P., Ridings, C. B., Mainer, L. L., Ganio, M. S., and Kavouras, S. A. (2014). Effect of air-filled vest on exercise-heat strain when wearing ballistic protection. *Ann. Occup. Hyg.* 58, 1057–1064. doi: 10.1093/annhyg/meu044

Almudehki, F., Girard, O., Grantham, J., and Racinais, S. (2012). Hot ambient conditions do not alter intermittent cycling sprint performance. *J. Sci. Med. Sport* 15, 148–152. doi: 10.1016/j.jsams.2011.07.009

Altrichter, S., Salow, J., Ardelean, E., Church, M. K., Werner, A., and Maurer, M. (2014). Development of a standardized pulse-controlled ergometry test for diagnosing and investigating cholinergic urticaria. *J. Dermatol. Sci.* 75, 88–93. doi: 10.1016/j.jdermsci.2014.04.007

Amano, T., Ichinose, M., Koga, S., Inoue, Y., Nishiyasu, T., and Kondo, N. (2011). Sweating responses and the muscle metaboreflex under mildly hyperthermic conditions in sprinters and distance runners. *J. Appl. Physiol.* 111, 524–529. doi: 10.1152/japplphysiol.00212.2011

Ando, S., Komiyama, T., Sudo, M., Kiyonaga, A., Tanaka, H., and Higaki, Y. (2015). The effects of temporal neck cooling on cognitive function during strenuous exercise in a hot environment: a pilot study. *BMC Res. Notes* 8:202. doi: 10.1186/s13104-015-1210-0

Arunachalam, K., Maccarini, P., Juang, T., Gaeta, C., and Stauffer, P. R. (2008). Performance evaluation of a conformal thermal monitoring sheet sensor array for measurement of surface temperature distributions during superficial hyperthermia treatments. *Int. J. Hyperth.* 24, 313–325. doi: 10.1080/02656730701881133

Bach, A. J. E., Stewart, I. B., Disher, A. E., and Costello, J. T. (2015a). A comparison between conductive and infrared devices for measuring mean skin temperature at rest, during exercise in the heat, and recovery. *PLoS ONE* 10:e0117907. doi: 10.1371/journal.pone.0117907

Bach, A. J. E., Stewart, I. B., Minett, G. M., and Costello, J. T. (2015b). Does the technique employed for skin temperature assessment alter outcomes? A systematic review. *Physiol. Meas.* 36, R27–R51. doi: 10.1088/0967-3334/36/9/R27

Bain, A. R., Deren, T. M., and Jay, O. (2011). Describing individual variation in local sweating during exercise in a temperate environment. *Eur. J. Appl. Physiol.* 111, 1599–1607. doi: 10.1007/s00421-010-1788-9

Bain, A. R., and Jay, O. (2011). Does summer in a humid continental climate elicit an acclimatization of human thermoregulatory responses? *Eur. J. Appl. Physiol.* 111, 1197–1205. doi: 10.1007/s00421-010-1743-9

Bandyopadhyay, A., Ping, F. W. C., and Keong, C. C. (2011). Effects of acute supplementation of caffeine and Panax ginseng on endurance running performance in a hot and humid environment. *J. Hum. Ergol. (Tokyo)* 40, 63–72.

Barr, D., Reilly, T., and Gregson, W. (2011). The impact of different cooling modalities on the physiological responses in firefighters during strenuous work performed in high environmental temperatures. *Eur. J. Appl. Physiol.* 111, 959–967. doi: 10.1007/s00421-010-1714-1

Barwood, M. J., Burrows, H., and Cessford, J. (2016). North Pole marathon laboratory lessons and field success. *Aerosp. Med. Hum. Perform.* 87, 493–497. doi: 10.3357/AMHP.4498.2016

Barwood, M. J., Corbett, J., Feeney, J., Hannaford, P., Henderson, D., Jones, I., et al. (2013). Compression garments: no enhancement of high-intensity exercise in hot radiant conditions. *Int. J. Sports Physiol. Perform.* 8, 527–535. doi: 10.1123/ijspp.8.5.527

Barwood, M. J., Corbett, J., Thomas, K., and Twentyman, P. (2015). Relieving thermal discomfort: effects of sprayed L-menthol on perception, performance, and time trial cycling in the heat. *Scand. J. Med. Sci. Sport.* 25, 211–218. doi: 10.1111/sms.12395

Barwood, M. J., Corbett, J., and White, D. K. (2014). Spraying with 0.20% L-menthol does not enhance 5 km running performance in the heat in untrained runners. *J. Sports Med. Phys. Fitness* 54, 595–604.

Barwood, M. J. M., Corbett, J., White, D., and James, J. (2012). Early change in thermal perception is not a driver of anticipatory exercise pacing in the heat. *Br. J. Sports Med.* 46, 936–942. doi: 10.1136/bjsports-2011-090536

Bedford, T., and Warner, C. G. (1934). On methods of measuring skin temperature. *J. Hyg.* 34, 81–98. doi: 10.1017/S0022172400034409

Bernard, V., Staffa, E., Mornstein, V., and Bourek, A. (2013). Infrared camera assessment of skin surface temperature - effect of emissivity. *Phys. Medica* 29, 583–591. doi: 10.1016/j.ejmp.2012.09.003

Best, S., Caillaud, C., and Thompson, M. (2012). The effect of ageing and fitness on thermoregulatory response to high-intensity exercise. *Scand. J. Med. Sci. Sports* 22, e29–e37. doi: 10.1111/j.1600-0838.2011.01384.x

Best, S., Thompson, M., Caillaud, C., Holvik, L., Fatseas, G., and Tammam, A. (2014). Exercise-heat acclimation in young and older trained cyclists. *J. Sci. Med. Sport* 17, 677–682. doi: 10.1016/j.jsams.2013.10.243

Blacker, S. D., Carter, J. M., Wilkinson, D. M., Richmond, V. L., Rayson, M. P., and Peattie, M. (2013). Physiological responses of police officers during job simulations wearing chemical, biological, radiological and nuclear personal protective equipment. *Ergonomics* 56, 137–147. doi: 10.1080/00140139.2012.734335

Bland, J. M., and Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 327, 307–310. doi: 10.1016/S0140-6736(86)90837-8

Boetcher, S. K. S., Sparrow, E. M., and Dugay, M. V. (2009). Characteristics of direct-contact, skin-surface temperature sensors. *Int. J. Heat Mass Transf.* 52, 3799–3804. doi: 10.1016/j.ijheatmasstransfer.2009.02.011

Bottoms, L., and Price, M. (2014). The effect of arm training on thermoregulatory responses and calf volume during upper body exercise. *Eur. J. Appl. Physiol.* 114, 1113–1122. doi: 10.1007/s00421-014-2842-9

Bourlai, T., Pryor, R. R., Suyama, J., Reis, S. E., and Hostler, D. (2012). Use of thermal imagery for estimation of core body temperature during precooling, exertion, and recovery in wildland firefighter protective clothing. *Prehospital Emerg. Care* 16, 390–399. doi: 10.3109/10903127.2012.670689

Brade, C., Dawson, B., and Wallman, K. (2014). Effects of different precooling techniques on repeat sprint ability in team sport athletes. *Eur. J. Sport Sci.* 14, S84–91. doi: 10.1080/17461391.2011.651491

Brown, J. J. S., and Connolly, D. A. (2015). Selected human physiological responses during extreme heat: the Badwater Ultramarathon. *J. Strength Cond. Res.* 29, 1729–1736. doi: 10.1519/JSC.0000000000000787

Bruning, R. S., Dahmus, J. D., Kenney, W. L., and Alexander, L. M. (2013). Aspirin and clopidogrel alter core temperature and skin blood flow during heat stress. *Med. Sci. Sports Exerc.* 45, 674–682. doi: 10.1249/MSS.0b013e31827981dc

Buono, M. J., and Ulrich, R. L. (1998). Comparison of mean skin temperature using "covered" versus "uncovered" contact thermistors. *Physiol. Meas.* 19, 297–300.

Burdon, C. A., Ruell, P., Johnson, N., Chapman, P., O'Brien, S., and O'Connor, H. T. (2015). The effect of ice-slushy consumption on plasma vasoactive intestinal peptide during prolonged exercise in the heat. *J. Therm. Biol.* 47, 59–62. doi: 10.1016/j.jtherbio.2014.11.005

Burton, A. C. (1935). Human calorimetry. II The average temperature of the tissues of the body. *J. Nutr.* 9, 261–280.

Byrne, C., Owen, C., Cosnefroy, A., and Lee, J. K. W. (2011). Self-paced exercise performance in the heat after pre-exercise cold-fluid ingestion. *J. Athl. Train.* 46, 592–599. doi: 10.4085/1062-6050-46.6.592

Caldwell, J. N., Patterson, M. J., and Taylor, N. A. S. (2012). Exertional thermal strain, protective clothing and auxiliary cooling In dry heat: evidence for physiological but not cognitive impairment. *Eur. J. Appl. Physiol.* 112, 3597–3606. doi: 10.1007/s00421-012-2340-x

Carlson, L. A., Ferguson, D. P., and Kenefick, R. W. (2014). Physiological strain of stock car drivers during competitive racing. *J. Therm. Biol.* 44, 20–26. doi: 10.1016/j.jtherbio.2014.06.001

Caruso, J. F., Barbosa, A., Erickson, L., Edwards, R., Perry, R., Learmonth, L., et al. (2015). Intermittent palm cooling's impact on resistive exercise performance. *Int. J. Sports Med.* 36, 814–821. doi: 10.1055/s-0035-1547264

Chen, Y., Lu, B., Chen, Y., and Feng, X. (2015). Breathable and stretchable temperature sensors inspired by skin. *Sci. Rep.* 5:11505. doi: 10.1038/srep11505

Childs, P. R. N. (2001). *Practical Temperature Measurement.* Oxford: Butterworth-Heinemann.

Childs, P. R. N., Greenwood, J. R., and Long, C. A. (2000). Review of temperature measurement. *Rev. Sci. Instrum.* 71, 2959–2978. doi: 10.1063/1.1305516

Choi, J. K., Miki, K., Sagawa, S., and Shiraki, K. (1997). Evaluation of mean skin temperature formulas by infrared thermography. *Int. J. Biometeorol.* 41, 68–75. doi: 10.1007/s004840050056

Chou, C., Tochihara, Y., Ismail, M. S., and Lee, J.-Y. (2011). Physiological strains of wearing aluminized and non-aluminized firefighters' protective clothing during exercise in radiant heat. *Ind. Health* 49, 185–194. doi: 10.2486/indhealth. MS1034

Colin, J., Timbal, J., Houdas, Y., Boutelier, C., and Guieu, J. D. (1971). Computation of mean body temperature from rectal and skin temperatures. *J. Appl. Physiol.* 31, 484–489. doi: 10.1152/jappl.1971.31.3.484

Coombs, G. B., Cramer, M. N., Ravanelli, N. M., Morris, N. B., and Jay, O. (2015). Acute acetaminophen ingestion does not alter core temperature or sweating during exercise in hot-humid conditions. *Scand. J. Med. Sci. Sport.* 25, 96–103. doi: 10.1111/sms.12336

Corbett, J., Barwood, M. J., and Tipton, M. J. (2015). Physiological cost and thermal envelope: a novel approach to cycle garment evaluation during a representative protocol. *Scand. J. Med. Sci. Sport.* 25, 152–158. doi: 10.1111/sms.12176

Coso, J. D., Hamouti, N., Ortega, J. F., Fernández-Elías, V. E., and Mora-Rodríguez, R. (2011). Relevance of individual characteristics for thermoregulation during exercise in a hot-dry environment. *Eur. J. Appl. Physiol.* 111, 2173–2181. doi: 10.1007/s00421-011-1847-x

Costello, J. T., Stewart, K. L., and Stewart, I. B. (2015). The effects of metabolic work rate and ambient environment on physiological tolerance times while wearing explosive and chemical personal protective equipment. *Biomed Res. Int.* 2015, 1–7. doi: 10.1155/2015/857536

Cotter, J. D., and Taylor, N. A. S. (2005). The distribution of cutaneous sudomotor and alliesthesial thermosensitivity in mildly heat-stressed humans: an open-loop approach. *J. Physiol.* 565, 335–345. doi: 10.1113/jphysiol.2004.081562

Crabtree, D. R., and Blannin, A. K. (2015). Effects of exercise in the cold on ghrelin, PYY, and food intake in overweight adults. *Med. Sci. Sports Exerc.* 47, 49–57. doi: 10.1249/MSS.0000000000000391

Cuddy, J. S., Buller, M., Hailes, W. S., and Ruby, B. C. (2013). Skin temperature and heart rate can be used to estimate physiological strain during exercise in the heat in a cohort of fit and unfit males. *Mil. Med.* 178, e841–e847. doi: 10.7205/MILMED-D-12-00524

Cuddy, J. S., Hailes, W. S., and Ruby, B. C. (2014). A reduced core to skin temperature gradient, not a critical core temperature, affects aerobic capacity in the heat. *J. Therm. Biol.* 43, 7–12. doi: 10.1016/j.jtherbio.2014.04.002

Cuddy, J. S., Sol, J. A., Hailes, W. S., and Ruby, B. C. (2015). Work patterns dictate energy demands and thermal strain during wildland firefighting. *Wilderness Environ. Med.* 26, 221–226. doi: 10.1016/j.wem.2014.12.010

Daanen, H. A. M. (2009). Manual performance deterioration in the cold estimated using the wind chill equivalent temperature. *Ind. Health* 47, 262–270. doi: 10.2486/indhealth.47.262

Daanen, H. A. M., Peerbooms, M., van den Hurk, C. J. G., van Os, B., Levels, K., Teunissen, L. P. J., et al. (2015). Core temperature affects scalp skin temperature during scalp cooling. *Int. J. Dermatol.* 54, 916–921. doi: 10.1111/ijd.12568

Davey, S. L., Barwood, M. J., and Tipton, M. J. (2013). Thermal perceptions and skin temperatures during continuous and intermittent ventilation of the torso throughout and after exercise in the heat. *Eur. J. Appl. Physiol.* 113, 2723–2735. doi: 10.1007/s00421-013-2697-5

Demachi, K., Yoshida, T., Kume, M., Tsuji, M., and Tsuneoka, H. (2013). The influence of internal and skin temperatures on active cutaneous vasodilation under different levels of exercise and ambient temperatures in humans. *Int. J. Biometeorol.* 57, 589–596. doi: 10.1007/s00484-012-0586-y

Demachi, K., Yoshida, T., Kume, M., and Tsuneoka, H. (2012a). The effect of starting or stopping skin cooling on the thermoregulatory responses during leg exercise in humans. *Int. J. Sports Med.* 33, 514–520. doi: 10.1055/s-0031-1301329

Demachi, K., Yoshida, T., and Tsuneoka, H. (2012b). Relationship between mean body temperature calculated by two- or three-compartment models and active cutaneous vasodilation in humans: a comparison between cool and warm environments during leg exercise. *Int. J. Biometeorol.* 56, 277–285. doi: 10.1007/s00484-011-0430-9

Deng, Z.-S., and Liu, J. (2008). "Effect of fixing material on skin-contact temperature measurement by wearable sensor," in *Proceedings of the 5th International Workshop on Wearable and Implantable Body Sensor Networks, BSN2008, in conjunction with the 5th International Summer School and Symposium on Medical Devices and Biosensors, ISSS-MDBS 2008* (Hong Kong: IEEE), 137–140.

De Pauw, K., De Geus, B., Roelands, B., Lauwens, F., Verschueren, J., Heyman, E., et al. (2011). Effect of five different recovery methods on repeated cycle performance. *Med. Sci. Sports Exerc.* 43, 890–897. doi: 10.1249/MSS.0b013e318200d25f

Deren, T. M., Coris, E. E., Bain, A. R., Walz, S. M., and Jay, O. (2012). Sweating is greater in NCAA football linemen independently of heat production. *Med. Sci. Sports Exerc.* 44, 244–252. doi: 10.1249/MSS.0b013e31822cb4e2

Deren, T. M., Coris, E. E., Casa, D. J., DeMartini, J. K., Bain, A. R., Walz, S. M., et al. (2014). Maximum heat loss potential is lower in football linemen during an NCAA summer training camp because of lower self-generated air flow. *J. Strength Cond. Res.* 28, 1656–1663. doi: 10.1519/JSC.0000000000000427

De Sousa, J., Cheatham, C., and Wittbrodt, M. (2014). The effects of a moisture-wicking fabric shirt on the physiological and perceptual responses during acute exercise in the heat. *Appl. Ergon.* 45, 1447–1453. doi: 10.1016/j.apergo.2014.04.006

Dollberg, S., Atherton, H. D., Sigda, M., Acree, C. M., and Hoath, S. B. (1994). Effect of insulated skin probes to increase skin-to-environmental temperature gradients of preterm infants cared for in convective incubators. *J. Pediatr.* 124, 799–801. doi: 10.1016/S0022-3476(05)81377-0

Eijsvogels, T. M. H., Bongers, C. C. W. G., Veltmeijer, M. T. W., Moen, M. H., and Hopman, M. (2014). Cooling during exercise in temperate conditions: impact on performance and thermoregulation. *Int. J. Sports Med.* 35, 840–846. doi: 10.1055/s-0034-1368723

Ely, B. R., Ely, M. R., and Cheuvront, S. N. (2011). Marginal effects of a large caffeine dose on heat balance during exercise-heat stress. *Int. J. Sport Nutr. Exerc. Metab.* 21, 65–70. doi: 10.1123/ijsnem.21.1.65

Faulkner, S. H., Hupperets, M., Hodder, S. G., and Havenith, G. (2015). Conductive and evaporative precooling lowers mean skin temperature and improves time trial performance in the heat. *Scand. J. Med. Sci. Sports* 25, 183–189. doi: 10.1111/sms.12373

Fernandes, A. A., Amorim, P. R., Brito, C. J., De Moura, A. G., Moreira, D. G., Costa, C. M. A., et al. (2014). Measuring skin temperature before, during and after exercise: a comparison of thermocouples and infrared thermography. *Physiol. Meas.* 35, 189–203. doi: 10.1088/0967-3334/35/2/189

Filho, E., Di Fronso, S., Mazzoni, C., Robazza, C., Bortoli, L., and Bertollo, M. (2015). My heart is racing! Psychophysiological dynamics of skilled racecar drivers. *J. Sports Sci.* 33, 945–959. doi: 10.1080/02640414.2014. 977940

Filingeri, D., Fournet, D., Hodder, S., and Havenith, G. (2015a). Mild evaporative cooling applied to the torso provides thermoregulatory benefits during running in the heat. *Scand. J. Med. Sci. Sports* 25, 200–210. doi: 10.1111/sms. 12322

Filingeri, D., Fournet, D., Hodder, S., and Havenith, G. (2015b). Tactile cues significantly modulate the perception of sweat-induced skin wetness independently of the level of physical skin wetness. *J. Neurophysiol.* 113, 3462–3473. doi: 10.1152/jn.00141.2015

Filingeri, D., Redortier, B., Hodder, S., and Havenith, G. (2014). Thermal and tactile interactions in the perception of local skin wetness at rest and during exercise in thermo-neutral and warm environments. *Neuroscience* 258, 121–130. doi: 10.1016/j.neuroscience.2013.11.019

Flesch, U., Wegener, O. H., Scheffler, A., and Ernst, H. (1976). Thermometry of the surface of human skin. A study on a model using thermocouples, thermistors, thermovision and thermodyes. *Phys. Med. Biol.* 21, 422–428. doi: 10.1088/0031-9155/21/3/008

Flouris, A. D., Dinas, P. C., Tsitoglou, K., Patramani, I., Koutedakis, Y., and Kenny, G. P. (2015). Non-invasive measurement of tibialis anterior muscle temperature during rest, cycling exercise and post-exercise recovery. *Physiol. Meas.* 36, N103–N113. doi: 10.1088/0967-3334/36/7/N103

Fonseca, S. F., Teles, M. C., Ribeiro, V. G. C., Magalhaes, F. C., Mendonca, V. A., Peixoto, M. F. D., et al. (2015). Hypertension is associated with greater heat exchange during exercise recovery in a hot environment. *Brazilian J. Med. Biol. Res.* 48, 1122–1129. doi: 10.1590/1414-431X20154532

Fortes, M. B., Di Felice, U., Dolci, A., Junglee, N. A., Crockford, M. J., West, L., et al. (2013). Muscle-damaging exercise increases heat strain during subsequent exercise heat stress. *Med. Sci. Sports Exerc.* 45, 1915–1924. doi: 10.1249/MSS.0b013e318294b0f8

Gagge, A. P., Stolwijk, J. A. J., and Hardy, J. D. (1967). Comfort and thermal sensations and associated physiological responses at various ambient temperatures. *Environ. Res.* 1, 1–20. doi: 10.1016/0013-9351(67)90002-3

Gagge, A. P., Winslow, C.-E. A., and Herrington, L. P. (1938). The influence of clothing on the physiological reactions of the human body to varying environmental temperatures. *Am. J. Physiol.* 124, 30–50. doi: 10.1152/ajplegacy.1938.124.1.30

Gagnon, D. D., Gagnon, S. S., Rintamäki, H., Törmäkangas, T., Puukka, K., Herzig, K.-H. H., et al. (2014a). The effects of cold exposure on leukocytes, hormones and cytokines during acute exercise in humans. *PLoS ONE* 9:e110774. doi: 10.1371/journal.pone.0110774

Gagnon, D. D., Rintamäki, H., Gagnon, S. S., Oksa, J., Porvari, K., Cheung, S. S., et al. (2014b). Fuel selection during short-term submaximal treadmill exercise in the cold is not affected by pre-exercise low-intensity shivering. *Appl. Physiol. Nutr. Metab.* 39, 282–291. doi: 10.1139/apnm-2013-0061

Gagnon, D., and Kenny, G. P. (2011). Sex modulates whole-body sudomotor thermosensitivity during exercise. *J. Physiol.* 589, 6205–6217. doi: 10.1113/jphysiol.2011.219220

Galloway, S. D., and Maughan, R. J. (1997). Effects of ambient temperature on the capacity to perform prolonged cycle exercise in man. *Med. Sci. Sports Exerc.* 29, 1240–1249. doi: 10.1097/00005768-199709000-00018

Gao, C., Kuklane, K., and Holmer, I. (2011). Cooling vests with phase change materials: the effects of melting temperature on heat strain alleviation in an extremely hot environment. *Eur. J. Appl. Physiol.* 111, 1207–1216. doi: 10.1007/s00421-010-1748-4

Gao, C., Lin, L.-Y., Halder, A., Kuklane, K., and Holmér, I. (2015). Validation of standard ASTM F2732 and comparison with ISO 11079 with respect to comfort temperature ratings for cold protective clothing. *Appl. Ergon.* 46, 44–53. doi: 10.1016/j.apergo.2014.07.002

Gao, W., Emaminejad, S., Nyein, H. Y. Y., Challa, S., Chen, K., Peck, A., et al. (2016). Fully integrated wearable sensor arrays for multiplexed *in situ* perspiration analysis. *Nature* 529, 509–514. doi: 10.1038/nature16521

Gillis, D. J., Weston, N., House, J. R., and Tipton, M. J. (2015). Influence of repeated daily menthol exposure on human temperature regulation and perception. *Physiol. Behav.* 139, 511–518. doi: 10.1016/j.physbeh.2014.12.009

Goh, S. S., Laursen, P. B., Dascombe, B., and Nosaka, K. (2011). Effect of lower body compression garments on submaximal and maximal running performance in cold (10°C) and hot (32°C) environments. *Eur. J. Appl. Physiol.* 111, 819–826. doi: 10.1007/s00421-010-1705-2

Griggs, K. E., Leicht, C. A., Price, M. J., and Goosey-Tolfrey, V. L. (2015). Thermoregulation during intermittent exercise in athletes with a spinal-cord injury. *Int. J. Sports Physiol. Perform.* 10, 469–475. doi: 10.1123/ijspp.2014-0361

Guadagni, D. N., Kreith, F., Smyth, C. J., and Bartholomew, B. A. (1972). Contact probe for skin temperature measurements. *J. Phys. E.* 5, 869–876. doi: 10.1088/0022-3735/5/9/012

Guéritée, J., House, J. R., Redortier, B., and Tipton, M. J. (2015a). The determinants of thermal comfort in cool water. *Scand. J. Med. Sci. Sports* 25, e459–e471. doi: 10.1111/sms.12360

Guéritée, J., Redortier, B., House, J. R., and Tipton, M. J. (2015b). Thermal comfort following immersion. *Physiol. Behav.* 139, 474–481. doi: 10.1016/j.physbeh.2014.12.016

Hardy, J. D. (1934a). The radiation of heat from the human body: I. An instrument for measuring the radiation and surface temperature of the skin. *J. Clin. Invest.* 13, 593–604. doi: 10.1172/JCI100607

Hardy, J. D. (1934b). The radiation of heat from the human body: II. A comparison of some methods of measurement. *J. Clin. Invest.* 13, 605–614. doi: 10.1172/JCI100608

Hardy, J. D., and Soderstrom, G. F. (1937). An improved apparatus for measuring surface and body temperature. *Rev. Sci. Instrum.* 8, 419–422. doi: 10.1063/1.1752205

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

Hayashi, K., Ito, N., Ichikawa, Y., and Suzuki, Y. (2014). Effect of postprandial thermogenesis on the cutaneous vasodilatory response during exercise. *Appl. Physiol. Nutr. Metab.* 39, 920–926. doi: 10.1139/apnm-2013-0506

Hayashi, K., Kawashima, T., and Suzuki, Y. (2012). Effect of menstrual cycle phase on the ventilatory response to rising body temperature during exercise. *J. Appl. Physiol.* 113, 237–245. doi: 10.1152/japplphysiol.01199.2011

Henriques, F. C. (1947). Studies of thermal injury. VIII. Automatic recording caloric applicator and skin-tissue and skin-surface thermocouples. *Rev. Sci. Instrum.* 18, 673–680. doi: 10.1063/1.1741041

Higgins, J. P. T., Altman, D. G., Sterne, J. A. C. (2011). "Chapter 8: Assessing risk of bias in included studies," in *Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0*, eds J. P. T. Higgins and S. Green (The Cochrane Collaboration). Available online at: www.cochrane-handbook.org

Hillman, A. R., Turner, M. C., Peart, D. J., Bray, J. W. J. W., Taylor, L., McNaughton, L. R. L. R., et al. (2013). A comparison of hyperhydration versus ad libitum fluid intake strategies on measures of oxidative stress, thermoregulation, and performance. *Res. Sport. Med.* 21, 305–317. doi: 10.1080/15438627.2013.825796

Hobson, R. M., Watson, P., and Maughan, R. J. (2013). Acute tryptophan depletion does not improve endurance cycling capacity in a warm environment. *Amino Acids* 44, 983–991. doi: 10.1007/s00726-012-1429-1

Holloway, G. A., Daly, C. H., Kennedy, D., and Chimoskey, J. (1976). Effects of external pressure loading on human skin blood flow measured by 133Xe clearance. *J. Appl. Physiol.* 40, 597–600. doi: 10.1152/jappl.1976.40.4.597

Ihsan, M., Watson, G., Lipski, M., and Abbiss, C. R. (2013). Influence of postexercise cooling on muscle oxygenation and blood volume changes. *Med. Sci. Sports Exerc.* 45, 876–882. doi: 10.1249/MSS.0b013e31827e13a2

International Organization for Standardization (2004). *ISO 9886:2004. Ergonomics—Evaluation of Thermal Strain by Physiological Measurements*. Geneva: International Organization for Standardization.

Ito, R., Nakano, M., Yamane, M., Amano, M., and Matsumoto, T. (2013). Effects of rain on energy metabolism while running in a cold environment. *Int. J. Sports Med.* 34, 707–711. doi: 10.1055/s-0032-1331770

Ito, R., Yamashita, N., Suzuki, E., and Matsumoto, T. (2015). Rain influences the physiological and metabolic responses to exercise in hot conditions. *J. Sports Sci.* 33, 892–898. doi: 10.1080/02640414.2014.977938

James, C. A., Richardson, A. J., Watt, P. W., and Maxwell, N. S. (2014). Reliability and validity of skin temperature measurement by telemetry thermistors and a thermal camera during exercise in the heat. *J. Therm. Biol.* 45, 141–149. doi: 10.1016/j.jtherbio.2014.08.010

Jay, O., DuCharme, M. B., Webb, P., Reardon, F. D., and Kenny, G. P. (2010). Estimating changes in volume-weighted mean body temperature using thermometry with an individualized correction factor. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 299, R387–R394. doi: 10.1152/ajpregu.00546.2009

Jay, O., Gariépy, L. M., Reardon, F. D., Webb, P., Ducharme, M. B., Ramsay, T., et al. (2007). A three-compartment thermometry model for the improved estimation of changes in body heat content. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 292, R167–R175. doi: 10.1152/ajpregu.00338.2006

Jirak, Z., Jokl, M., Stverak, J., Pechlat, R., and Coufalov, H. (1975). Correction factors in skin temperature measurement. *J. Appl. Physiol.* 38, 752–756. doi: 10.1152/jappl.1975.38.4.752

Jovanović, D., Karkalić, R., Zeba, S., Pavlović, M., and Radaković, S. S. (2014). Physiological tolerance to uncompensated heat stress in soldiers: effects of various types of body cooling systems. *Vojnosanit. Pregl.* 71, 259–264. doi: 10.2298/VSP120731045J

Jutte, L. S., Knight, K. L., Long, B. C., Hawkins, J. R., Schulthies, S. S., and Dalley, E. B. (2005). The uncertainty (validity and reliability) of three electrothermometers in therapeutic modality research. *J. Athl. Train.* 40, 207–210.

Katica, C. P., Pritchett, R. C., Pritchett, K. L., Del Pozzi, A. T., Balilionis, G., and Burnham, T. (2011). Effects of forearm vs. leg submersion in work tolerance time in a hot environment while wearing firefighter protective clothing. *J. Occup. Environ. Hyg.* 8, 473–477. doi: 10.1080/15459624.2011.590743

Kaufman, W. C., and Pittman, J. C. (1966). Quantitative radiometric measurement of skin temperature. *J. Appl. Physiol.* 21, 302–304. doi: 10.1152/jappl.1966.21.1.302

Kenefick, R. W., Cheuvront, S. N., Ely, B. R., Palombo, L. J., and Sawka, M. N. (2011). DEET insect repellent: effects on thermoregulatory sweating and physiological strain. *Eur. J. Appl. Physiol.* 111, 3061–3068. doi: 10.1007/s00421-011-1932-1

Kenefick, R. W., Sollanek, K. J., Charkoudian, N., and Sawka, M. N. (2014). Impact of skin temperature and hydration on plasma volume responses during exercise. *J. Appl. Physiol.* 117, 413–420. doi: 10.1152/japplphysiol.00415.2014

Keramidas, M. E., Geladas, N. D., Mekjavic, I. B., and Kounalakis, S. N. (2013). Forearm-finger skin temperature gradient as an index of cutaneous perfusion during steady-state exercise. *Clin. Physiol. Funct. Imaging* 33, 400–404. doi: 10.1111/cpf.12043

Klute, G. K., Huff, E., and Ledoux, W. R. (2014). Does activity affect residual limb skin temperatures? *Clin. Orthop. Relat. Res.* 472, 3062–3067. doi: 10.1007/s11999-014-3741-4

Kounalakis, S. N., Eiken, O., and Mekjavic, I. B. (2012). Exercise thermoregulatory responses following a 28-day sleep-high train-low regimen. *Eur. J. Appl. Physiol.* 112, 3881–3891. doi: 10.1007/s00421-012-2374-0

Krause, B. F. (1993). Accuracy and response time comparisons of four skin temperature-monitoring devices. *Nurse Anesth.* 4, 55–61.

Lee, E. R., Kapp, D. S., Lohrbach, A. W., and Sokol, J. L. (1994). Influence of water bolus temperature on measured skin surface and intradermal temperatures. *Int. J. Hyperth.* 10, 59–72. doi: 10.3109/02656739409009332

Lee, J. F., Christmas, K. M., Machin, D. R., McLean, B. D., and Coyle, E. F. (2015). Warm skin alters cardiovascular responses to cycling after preheating and precooling. *Med. Sci. Sports Exerc.* 47, 1168–1176. doi: 10.1249/MSS.0000000000000539

Lee, J.-Y., Yamamoto, Y., Oe, R., Son, S.-Y., Wakabayashi, H., and Tochihara, Y. (2014a). The European, Japanese and US protective helmet, gloves and boots for firefighters: thermoregulatory and psychological evaluations. *Ergonomics* 57, 1213–1221. doi: 10.1080/00140139.2014.914578

Lee, J. K. W., Koh, A. C. H., Koh, S. X. T., Liu, G. J. X., Nio, A. Q. X., and Fan, P. W. P. (2014b). Neck cooling and cognitive performance following exercise-induced hyperthermia. *Eur. J. Appl. Physiol.* 114, 375–384. doi: 10.1007/s00421-013-2774-9

Lee, J. K. W., Yeo, Z. W., Nio, A. Q. X., Koh, A. C. H., Teo, Y. S., Goh, L. F., et al. (2013). Cold drink attenuates heat strain during work-rest cycles. *Int. J. Sports Med.* 34, 1037–1042. doi: 10.1055/s-0033-1337906

Lee, J.-Y. Y., Nakao, K., Bakri, I., and Tochihara, Y. (2012). Body regional influences of L-menthol application on the alleviation of heat strain while wearing firefighter's protective clothing. *Eur. J. Appl. Physiol.* 112, 2171–2183. doi: 10.1007/s00421-011-2192-9

Lenhardt, R., and Sessler, D. I. (2006). Estimation of mean-body temperature from mean-skin and core temperature. *Anesthesiology* 105, 1117–1121. doi: 10.1097/00000542-200612000-00011

Levels, K., de Koning, J., Broekhuijzen, I., Zwaan, T., Foster, C., and Daanen, H. A. M. (2014a). Effects of radiant heat exposure on pacing pattern during a 15-km cycling time trial. *J. Sports Sci.* 32, 845–852. doi: 10.1080/02640414.2013.862843

Levels, K., de Koning, J. J., Foster, C., and Daanen, H. A. M. (2012). The effect of skin temperature on performance during a 7.5-km cycling time trial. *Eur. J. Appl. Physiol.* 112, 3387–3395. doi: 10.1007/s00421-012-2316-x

Levels, K., de Koning, J. J., Mol, E., Foster, C., and Daanen, H. A. M. (2014b). The effect of pre-warming on performance during simulated firefighting exercise. *Appl. Ergon.* 45, 1504–1509. doi: 10.1016/j.apergo.2014.04.011

Levels, K., Teunissen, L., P., J., de Haan, A., de Koning, J. J., van Os, B., and Daanen, H., A. M. (2013). Effect of warm-up and precooling on pacing during a 15-km cycling time trial in the heat. *Int. J. Sports Physiol. Perform.* 8, 307–311. doi: 10.1123/ijspp.8.3.307

Libert, J. P., Candas, V., and Vogt, J. J. (1978). Sweating response in man during transient rises of air temperature. *J. Appl. Physiol.* 44, 284–290. doi: 10.1152/jappl.1978.44.2.284

Limbaugh, J. D., Wimer, G. S., Long, L. H., and Baird, W. H. (2013). Body fatness, body core temperature, and heat loss during moderate-intensity exercise. *Aviat. Space Environ. Med.* 84, 1153–1158. doi: 10.3357/ASEM.3627.2013

Liu, Y., Zhu, S. H., Wang, G. H., Ye, F., and Li, P. Z. (2013). Validity and reliability of multiparameter physiological measurements recorded by the Equivital LifeMonitor during activities of various intensities. *J. Occup. Environ. Hyg.* 10, 78–85. doi: 10.1080/15459624.2012.747404

Livingstone, S. D., Nolan, R. W., Frim, J., Reed, L. D., and Limmer, R. E. (1987). A thermographic study of the effect of body composition and ambient temperature on the accuracy of mean skin temperature calculations. *Eur. J. Appl. Physiol. Occup. Physiol.* 56, 120–125. doi: 10.1007/BF00696387

MacRae, B. A., Laing, R. M., Niven, B. E., and Cotter, J. D. (2012). Pressure and coverage effects of sporting compression garments on cardiovascular function, thermoregulatory function, and exercise performance. *Eur. J. Appl. Physiol.* 112, 1783–1795. doi: 10.1007/s00421-011-2146-2

MacRae, B. A., Laing, R. M., Wilson, C. A., and Niven, B. E. (2014). Layering garments during rest and exercise in the cold (8°C): wearer responses and comparability with selected fabric properties. *Ergonomics* 57, 271–281. doi: 10.1080/00140139.2013.869356

Mahanty, S. D., and Roemer, R. B. (1979a). Skin temperature probe. *J. Biomech. Eng.* 101, 232–238. doi: 10.1115/1.3426251

Mahanty, S. D., and Roemer, R. B. (1979b). The effect of pressure on skin temperature measurements for a disk sensor. *J. Biomech. Eng.* 101, 261–266. doi: 10.1115/1.3426255

Malone, E. W. (1963). Surface temperature measurement errors. *IEEE Trans. Aerosp.* 1, 15–20. doi: 10.1109/TA.1963.4319342

Martinez, N., Psikuta, A., Kuklane, K., Quesada, J. I. P., de Anda, R. M. C. O., Soriano, P. P., et al. (2016). Validation of the thermophysiological model by Fiala for prediction of local skin temperatures. *Int. J. Biometeorol.* 60, 1969–1982. doi: 10.1007/s00484-016-1184-1

Matsuzuki, H., Ito, A., Ayabe, M., Haruyama, Y., Tomita, S., Katamoto, S., et al. (2011). The effects of work environments on thermal strain on workers in commercial kitchens. *Ind. Health* 49, 605–613. doi: 10.2486/indhealth.MS1219

Mauger, A. R., Taylor, L., Harding, C., Wright, B., Foster, J., and Castle, P. C. (2014). Acute acetaminophen (paracetamol) ingestion improves time to exhaustion during exercise in the heat. *Exp. Physiol.* 99, 164–171. doi: 10.1113/expphysiol.2013.075275

Maughan, R. J., Otani, H., and Watson, P. (2012). Influence of relative humidity on prolonged exercise capacity in a warm environment. *Eur. J. Appl. Physiol.* 112, 2313–2321. doi: 10.1007/s00421-011-2206-7

McFarlin, B. K., Venable, A. S., Williams, R. R., and Jackson, A. W. (2015). Comparison of techniques for the measurement of skin temperature during exercise in a hot, humid environment. *Biol. Sport* 32, 11–14. doi: 10.5604/20831862.1124569

McGlone, B., and Bazett, H. C. (1927). The temperature of the air in contact with the skin. *Am. J. Physiol.* 82, 452–461. doi: 10.1152/ajplegacy.1927.82.2.452

McLellan, T. M., Daanen, H. A. M., and Cheung, S. S. (2013). Encapsulated environment. *Compr. Physiol.* 3, 1363–1391. doi: 10.1002/cphy.c130002

Mee, J. A., Doust, J., and Maxwell, N. S. (2015). Repeatability of a running heat tolerance test. *J. Therm. Biol.* 49–50, 91–97. doi: 10.1016/j.jtherbio.2015.02.010

Michalski, L., Eckersdorf, K., Kucharski, J., and McGhee, J. (2001). *Temperature Measureoment*, 2nd Edn. Chichester, UK: John Wiley & Sons, Ltd.

Mieras, M. E., Heesch, M. W. S., and Slivka, D. R. (2014). Physiological and psychological responses to outdoor vs. laboratory cycling. *J. Strength Cond. Res.* 28, 2324–2329. doi: 10.1519/JSC.0000000000000384

Miyagawa, K., Kamijo, Y.-I., Ikegawa, S., Goto, M., and Nose, H. (2011). Reduced hyperthermia-induced cutaneous vasodilation and enhanced exercise-induced plasma water loss at simulated high altitude (3,200 m) in humans. *J. Appl. Physiol.* 110, 157–165. doi: 10.1152/japplphysiol.00950.2010

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., and PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann. Intern. Med.* 151, 264–269. doi: 10.7326/0003-4819-151-4-200908180-00135

Moore, J. P., Harper Smith, A. D., Di Felice, U., and Walsh, N. P. (2013). Three nights of sleep deprivation does not alter thermal strain during exercise in the heat. *Eur. J. Appl. Physiol.* 113, 2353–2360. doi: 10.1007/s00421-013-2671-2

Moran, D. S., Castellani, J. W., O'Brien, C., Young, A. J., and Pandolf, K. B. (1999). Evaluating physiological strain during cold exposure using a new cold strain index. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 277, R556–R564.

Mora-Rodriguez, R., Aguado-Jimenez, R., Del Coso, J., and Estevez, E. (2012). A standard blood bank donation alters the thermal and cardiovascular responses during subsequent exercise. *Transfusion* 52, 2339–2347. doi: 10.1111/j.1537-2995.2012.03613.x

Mora-Rodriguez, R., Ortega, J. F., and Hamouti, N. (2011). In a hot-dry environment racewalking increases the risk of hyperthermia in comparison to when running at a similar velocity. *Eur. J. Appl. Physiol.* 111, 1073–1080. doi: 10.1007/s00421-010-1733-y

Morris, N. B., Bain, A. R., Cramer, M. N., and Jay, O. (2014). Evidence that transient changes in sudomotor output with cold and warm fluid ingestion are

independently modulated by abdominal, but not oral thermoreceptors. *J. Appl. Physiol.* 116, 1088–1095. doi: 10.1152/japplphysiol.01059.2013

Morrison, S. A., Cheung, S., and Cotter, J. D. (2014a). Importance of airflow for physiologic and ergogenic effects of precooling. *J. Athl. Train.* 49, 632–639. doi: 10.4085/1062-6050-49.3.27

Morrison, S. A., Cheung, S. S., and Cotter, J. D. (2014b). Bovine colostrum, training status, and gastrointestinal permeability during exercise in the heat: a placebo-controlled double-blind study. *Appl. Physiol. Nutr. Metab.* 39, 1070–1082. doi: 10.1139/apnm-2013-0583

Moyen, N. E., Ellis, C. L. V., Ciccone, A. B., Thurston, T. S., Cochrane, K. C., Brown, L. E., et al. (2014a). Increasing relative humidity impacts low-intensity exercise in the heat. *Aviat. Space Environ. Med.* 85, 112–119. doi: 10.3357/ASEM.3787.2014

Moyen, N. E., Mundel, T., Du Bois, A. M., Ciccone, A. B., Morton, R. H., and Judelson, D. A. (2014b). Increasing humidity affects thermoregulation during low-intensity exercise in women. *Aviat. Space Environ. Med.* 85, 905–911. doi: 10.3357/ASEM.3993.2014

Muller, M. D., Muller, S. M., Ryan, E. J., Bellar, D. M., Kim, C.-H., and Glickman, E. L. (2011). Pain and thermal sensation in the cold: the effect of interval versus continuous exercise. *Eur. J. Appl. Physiol.* 111, 979–987. doi: 10.1007/s00421-010-1726-x

Muñoz, C. X., Carney, K. R., Schick, M. K., Coburn, J. W., Becker, A. J., Judelson, D. A., et al. (2012). Effects of oral rehydration and external cooling on physiology, perception, and performance in hot, dry climates. *Scand. J. Med. Sci. Sport.* 22, e115–e124. doi: 10.1111/j.1600-0838.2012.01510.x

Nadel, E. R., Bullard, R. W., and Stolwijk, J. A. J. (1971). Importance of skin temperature in the regulation of sweating. *J. Appl. Physiol.* 31, 80–87. doi: 10.1152/jappl.1971.31.1.80

Nassif, C., Gomes, A. R., Peixoto, G. H. C., Chagas, M. H., Soares, D. D., Silami-Garcia, E., et al. (2014). The effect of double-Blind carbohydrate ingestion during 60 km of self-paced exercise in warm ambient conditions. *PLoS ONE* 9:e104710. doi: 10.1371/journal.pone.0104710

Nicholas, J. V., and White, D. R. (2001). *Traceable Temperatures*. Chichester, UK: John Wiley & Sons, Ltd.

Niedermann, R., Wyss, E., Annaheim, S., Psikuta, A., Davey, S., and Rossi, R. M. (2014). Prediction of human core body temperature using non-invasive measurement methods. *Int. J. Biometeorol.* 58, 7–15. doi: 10.1007/s00484-013-0687-2

Nielsen, R., Olesen, B. W., and Fanger, P. O. (1985). Effect of physical activity and air velocity on the thermal insulation of clothing. *Ergonomics* 28, 1617–1631. doi: 10.1080/00140138508963299

Noonan, B., and Stachenfeld, N. (2012). The effects of undergarment composition worn beneath hockey protective equipment on high-intensity intermittent exercise. *J. Strength Cond. Res.* 26, 2309–2316. doi: 10.1519/JSC.0b013e3182606e1e

Oda, S., and Shirakawa, K. (2014). Sleep onset is disrupted following pre-sleep exercise that causes large physiological excitement at bedtime. *Eur. J. Appl. Physiol.* 114, 1789–1799. doi: 10.1007/s00421-014-2873-2

Oksa, J., Hosio, S., Makinen, T., Lindholm, H., Rintamaki, H., Rissanen, S., et al. (2014). Muscular, cardiorespiratory and thermal strain of mast and pole workers. *Ergonomics* 57, 669–678. doi: 10.1080/00140139.2014.895854

Parker, S. M., Erin, J. R., Pryor, R. R., Khorana, P., Suyama, J., Guyette, F. X., et al. (2013). The effect of prolonged light intensity exercise in the heat on executive function. *Wilderness Environ. Med.* 24, 203–210. doi: 10.1016/j.wem.2013.01.010

Pearson, J., Low, D. A., Stohr, E., Kalsi, K., Ali, L., Barker, H., et al. (2011). Hemodynamic responses to heat stress in the resting and exercising human leg: insight into the effect of temperature on skeletal muscle blood flow. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 300, R663–R673. doi: 10.1152/ajpregu.00662.2010

Pennes, H. H. (1948). Analysis of tissue and arterial blood temperatures in the resting human forearm. *J. Appl. Physiol.* 1, 93–122. doi: 10.1152/jappl.1948.1.2.93

Pèriard, J. D., Cramer, M. N., Chapman, P. G., Caillaud, C., and Thompson, M. W. (2011). Cardiovascular strain impairs prolonged self-paced exercise in the heat. *Exp. Physiol.* 96, 134–144. doi: 10.1113/expphysiol.2010.054213

Pèriard, J. D., and Racinais, S. (2015). Heat stress exacerbates the reduction in middle cerebral artery blood velocity during prolonged self-paced exercise. *Scand. J. Med. Sci. Sport.* 25, 135–144. doi: 10.1111/sms.12379

Pèriard, J. D., Racinais, S. S., Knez, W. L., Herrera, C. P., Christian, R. J., Girard, O., et al. (2014). Thermal, physiological and perceptual strain mediate alterations in match-play tennis under heat stress. *Br. J. Sports Med.* 48, i32–i38. doi: 10.1136/bjsports-2013-093063

Pèriard, J. D., Thompson, M. W., Caillaud, C., and Quaresima, V. (2013). Influence of heat stress and exercise intensity on vastus lateralis muscle and prefrontal cortex oxygenation. *Eur. J. Appl. Physiol.* 113, 211–222. doi: 10.1007/s00421-012-2427-4

Ping, F. W. C., Keong, C. C., and Bandyopadhyay, A. (2011). Effects of acute supplementation of Panax ginseng on endurance running in a hot and humid environment. *Indian J. Med. Res.* 133, 96–102.

Priego Quesada, J. I., Martínez Guillamón, N., de Anda, R. M. C. O., Psikuta, A., Annaheim, S., Rossi, R. M., et al. (2015). Effect of perspiration on skin temperature measurements by infrared thermography and contact thermometry during aerobic cycling. *Infrared Phys. Technol.* 72, 68–76. doi: 10.1016/j.infrared.2015.07.008

Pryor, R. R., Suyama, J., Guyette, F. X., Reis, S. E., and Hostler, D. (2015). The effects of ice slurry ingestion before exertion in Wildland firefighting gear. *Prehospital Emerg. Care* 19, 241–246. doi: 10.3109/10903127.2014.959221

Psikuta, A., Niedermann, R., and Rossi, R. M. (2014). Effect of ambient temperature and attachment method on surface temperature measurements. *Int. J. Biometeorol.* 58, 877–885. doi: 10.1007/s00484-013-0669-4

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

Randall, C. A., Ross, E. Z., and Maxwell, N. S. (2015). Effect of practical precooling on neuromuscular function and 5-km time-trial performance in hot, humid conditions among well-trained male runners. *J. Strength Cond. Res.* 29, 1925–1936. doi: 10.1519/JSC.0000000000000840

Ravanelli, N. M., Cramer, M. N., Molgat-Seon, Y., Carlsen, A. N., and Jay, O. (2014). Do greater rates of body heat storage precede the accelerated reduction of self-paced exercise intensity in the heat? *Eur. J. Appl. Physiol.* 114, 2399–2410. doi: 10.1007/s00421-014-2958-y

Regan, J. M., Macfarlane, D. J., and Taylor, N. A. (1996). An evaluation of the role of skin temperature during heat adaptation. *Acta Physiol. Scandanavia* 158, 365–375. doi: 10.1046/j.1365-201X.1996.561311000.x

Richmond, V. L., Davey, S., Griggs, K., and Havenith, G. (2015). Prediction of core body temperature from multiple variables. *Ann. Occup. Hyg.* 59, 1168–1178. doi: 10.1093/annhyg/mev054

Rintamäki, H., Kyröläinen, H., Santtila, M., Mäntysaari, M., Simonen, R., Torpo, H., et al. (2012). From the subarctic to the tropics: effects of 4-month deployment on soldiers' heat stress, heat strain, and physical performance. *J. Strength Cond. Res.* 26, 45–52. doi: 10.1519/JSC.0b013e31825d817e

Roberge, R., Benson, S., and Kim, J.-H. (2012a). Thermal burden of N95 filtering facepiece respirators. *Ann. Occup. Hyg.* 56, 808–814. doi: 10.1093/annhyg/mes001

Roberge, R. J., Kim, J.-H. H., and Benson, S. M. (2012b). Absence of consequential changes in physiological, thermal and subjective responses from wearing a surgical mask. *Respir. Physiol. Neurobiol.* 181, 29–35. doi: 10.1016/j.resp.2012.01.010

Ross, M., Abbiss, C., Laursen, P., Martin, D., and Burke, L. (2013). Precooling methods and their effects on athletic performance: a systematic review and practical applications. *Sport. Med.* 43, 207–225. doi: 10.1007/s40279-012-0014-9

Rossi, R. (2003). Fire fighting and its influence on the body. *Ergonomics* 46, 1017–1033. doi: 10.1080/0014013031000121968

Rowell, L. B., Murray, J. A., Brengelmann, G. L., and Kraning, K. K. (1969). Human cardiovascular adjustments to rapid changes in skin temperature during exercise. *Circ. Res.* 24, 711–724. doi: 10.1161/01.RES.24.5.711

Ruckman, J. E., Murray, R., and Choi, H. S. (1999). Engineering of clothing systems for improved thermophysiological comfort: the effect of openings. *Int. J. Cloth. Sci. Technol.* 11, 37–52. doi: 10.1108/09556229910258098

Sandsund, M., Saursaunet, V., Wiggen, Ø., Renberg, J., Færevik, H., and van Beekvelt, M. C. P. (2012). Effect of ambient temperature on endurance

performance while wearing cross-country skiing clothing. *Eur. J. Appl. Physiol.* 112, 3939–3947. doi: 10.1007/s00421-012-2373-1

Sawka, M. N., Cheuvront, S. N., and Kenefick, R. W. (2012). High skin temperature and hypohydration impair aerobic performance. *Exp. Physiol.* 97, 327–332. doi: 10.1113/expphysiol.2011.061002

Sawka, M. N., Leon, L. R., Montain, S. J., and Sonna, L. A. (2011). Integrated physiological mechanisms of exercise performance, adaptation, and maladaptation to heat stress. *Compr. Physiol.* 1, 1883–1928. doi: 10.1002/cphy.c100082

Schlader, Z. J., Simmons, S. E., Stannard, S. R., and Mundel, T. (2011a). Skin temperature as a thermal controller of exercise intensity. *Eur. J. Appl. Physiol.* 111, 1631–1639. doi: 10.1007/s00421-010-1791-1

Schlader, Z. J., Simmons, S. E., Stannard, S. R., and Mundel, T. (2011b). The independent roles of temperature and thermal perception in the control of human thermoregulatory behavior. *Physiol. Behav.* 103, 217–224. doi: 10.1016/j.physbeh.2011.02.002

Schlader, Z. J., Stannard, S. R., and Mundel, T. (2011c). Is peak oxygen uptake a determinant of moderate-duration self-paced exercise performance in the heat? *Appl. Physiol. Nutr. Metab.* 36, 863–872. doi: 10.1139/h11-111

Shimazaki, Y., and Murata, M. (2015). Effect of gait on formation of thermal environment inside footwear. *Appl. Ergon.* 49, 55–62. doi: 10.1016/j.apergo.2015.01.007

Siegel, R., Mate, J., Watson, G., Nosaka, K., and Laursen, P. B. (2011). The influence of ice slurry ingestion on maximal voluntary contraction following exercise-induced hyperthermia. *Eur. J. Appl. Physiol.* 111, 2517–2524. doi: 10.1007/s00421-011-1876-5

Sinclair, W. H., and Brownsberger, J. C. (2013). Wearing long pants while working outdoors in the tropics does not yield higher body temperatures. *Aust. N. Z. J. Public Health* 37, 70–75. doi: 10.1111/1753-6405.12013

Skein, M., Duffield, R., Cannon, J., and Marino, F. E. (2012). Self-paced intermittent-sprint performance and pacing strategies following respective pre-cooling and heating. *Eur. J. Appl. Physiol.* 112, 253–266. doi: 10.1007/s00421-011-1972-6

Smith, D. L., Haller, J. M., Hultquist, E. M., Lefferts, W. K., and Fehling, P. C. (2013). Effect of clothing layers in combination with fire fighting personal protective clothing on physiological and perceptual responses to intermittent work and on materials performance test results. *J. Occup. Environ. Hyg.* 10, 259–269. doi: 10.1080/15459624.2013.769841

Sperlich, B., and Holmberg, H.-C. (2017). Wearable, yes, but able …?: it is time for evidence-based marketing claims! *Br. J. Sports Med.* 51:1240. doi: 10.1136/bjsports-2016-097295

Spitz, M. G., Kenefick, R. W., and Mitchell, J. B. (2014). The effects of elapsed time after warm-up on subsequent exercise performance in a cold environment. *J. Strength Cond. Res.* 28, 1351–1357. doi: 10.1519/JSC.0000000000000291

Stapleton, J. M., Yardley, J. E., Boulay, P., Sigal, R. J., and Kenny, G. P. (2013). Whole-body heat loss during exercise in the heat is not impaired in type 1 diabetes. *Med. Sci. Sports Exerc.* 45, 1656–1664. doi: 10.1249/MSS.0b013e31829002f3

Steketee, J. (1973). Spectral emissivity of skin and pericardium. *Phys. Med. Biol.* 18, 686–694. doi: 10.1088/0031-9155/18/5/307

Stevens, C. J., Taylor, L., and Dascombe, B. J. (2017). Cooling during exercise: an overlooked strategy for enhancing endurance performance in the heat. *Sport. Med.* 47, 829–841. doi: 10.1007/s40279-016-0625-7

Sun, C., Au, J. S., chuen, Fan, J., and Zheng, R. (2015). Novel ventilation design of combining spacer and mesh structure in sports T-shirt significantly improves thermal comfort. *Appl. Ergon.* 48, 138–147. doi: 10.1016/j.apergo.2014.11.012

Suzuki, E., Ohya, T., Ito, R., Matsumoto, T., and Kitagawa, K. (2014). Physiological responses in alpine skiers during on-snow training simulation in the cold. *Int. J. Sports Med.* 35, 392–398. doi: 10.1055/s-0033-1353146

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

Temfemo, A., Carling, C., and Ahmaidi, S. (2011). Relationship between power output, lactate, skin temperature, and muscle activity during brief repeated exercises with increasing intensity. *J. Strength Cond. Res.* 25, 915–921. doi: 10.1519/JSC.0b013e3181d680f0

Teunissen, L. P. J., De Haan, A., De Koning, J. J., and Daanen, H. A. M. (2013). Effects of wind application on thermal perception and self-paced performance. *Eur. J. Appl. Physiol.* 113, 1705–1717. doi: 10.1007/s00421-013-2596-9

Teunissen, L. P. J., Wang, L. C. L.-C., Chou, S.-N. S. N., Huang, C.-H. C., hsien, Jou, G.-T. G. T., and Daanen, H. A. M. (2014). Evaluation of two cooling systems under a firefighter coverall. *Appl. Ergon.* 45, 1433–1438. doi: 10.1016/j.apergo.2014.04.008

Tokizawa, K., Sawada, S., Oka, T., Yasuda, A., Tai, T., Ida, H., et al. (2014). Fan-precooling effect on heat strain while wearing protective clothing. *Int. J. Biometeorol.* 58, 1919–1925. doi: 10.1007/s00484-014-0794-8

Tsuji, B., Honda, Y., Fujii, N., Kondo, N., and Nishiyasu, T. (2012). Comparison of hyperthermic hyperventilation during passive heating and prolonged light and moderate exercise in the heat. *J. Appl. Physiol.* 113, 1388–1397. doi: 10.1152/japplphysiol.00335.2012

Tsuji, M., Kume, M., Tuneoka, H., and Yoshida, T. (2014). Differences in the heat stress associated with white sportswear and being semi-nude in exercising humans under conditions of radiant heat and wind at a wet bulb globe temperature of greater than 28°C. *Int. J. Biometeorol.* 58, 1393–1402. doi: 10.1007/s00484-013-0722-3

Tucker, R., Marle, T., Lambert, E. V., and Noakes, T. D. (2006). The rate of heat storage mediates an anticipatory reduction in exercise intensity during cycling at a fixed rating of perceived exertion. *J. Physiol.* 574, 905–915. doi: 10.1113/jphysiol.2005.101733

Tumilty, L., Davison, G., Beckmann, M., and Thatcher, R. (2011). Oral tyrosine supplementation improves exercise capacity in the heat. *Eur. J. Appl. Physiol.* 111, 2941–2950. doi: 10.1007/s00421-011-1921-4

Tumilty, L., Davison, G., Beckmann, M., and Thatcher, R. (2014). Failure of oral tyrosine supplementation to improve exercise performance in the heat. *Med. Sci. Sports Exerc.* 46, 1417–1425. doi: 10.1249/MSS.0000000000000243

Turner, A. P., and Richards, H. (2015). Physiological and selective attention demands during an international rally motor sport event. *Biomed Res. Int.* 2015:638659. doi: 10.1155/2015/638659

Tyler, C. J. (2011). The effect of skin thermistor fixation method on weighted mean skin temperature. *Physiol. Meas.* 32, 1541–1547. doi: 10.1088/0967-3334/32/10/003

Tyler, C. J., and Sunderland, C. (2011a). Cooling the neck region during exercise in the heat. *J. Athl. Train.* 46, 61–68. doi: 10.4085/1062-6050-46.1.61

Tyler, C. J., and Sunderland, C. (2011b). Neck cooling and running performance in the heat: single versus repeated application. *Med. Sci. Sports Exerc.* 43, 2388–2395. doi: 10.1249/MSS.0b013e318222ef72

Van Cutsem, J., Pattyn, N., Vissenaeken, D., Dhondt, G., De Pauw, K., Tonoli, C., et al. (2015). The influence of a mild thermal challenge and severe hypoxia on exercise performance and serum BDNF. *Eur. J. Appl. Physiol.* 115, 2135–2148. doi: 10.1007/s00421-015-3193-x

van Marken Lichtenbelt, W. D., Daanen, H. A. M., Wouters, L., Fronczek, R., Raymann, R. J. E. M., Severens, N. M. W., et al. (2006). Evaluation of wireless determination of skin temperature using iButtons. *Physiol. Behav.* 88, 489–497. doi: 10.1016/j.physbeh.2006.04.026

Vanos, J. K., Warland, J. S., Gillespie, T. J., and Kenny, N. A. (2012). Thermal comfort modelling of body temperature and psychological variations of a human exercising in an outdoor environment. *Int. J. Biometeorol.* 56, 21–32. doi: 10.1007/s00484-010-0393-2

Veltmeijer, M. T., Pluim, B., Thijssen, D. H., Hopman, M. T., and Eijsvogels, T. M. (2014). Thermoregulatory responses in wheelchair tennis players: a pilot study. *Spinal Cord* 52, 373–377. doi: 10.1038/sc.2014.27

Wakabayashi, H., Wijayanto, T., Lee, J.-Y., Hashiguchi, N., Saat, M., and Tochihara, Y. (2011). Comparison of heat dissipation response between Malaysian and Japanese males during exercise in humid heat stress. *Int. J. Biometeorol.* 55, 509–517. doi: 10.1007/s00484-010-0374-5

Wakabayashi, H., Wijayanto, T., Lee, J.-Y., Hashiguchi, N., Saat, M., and Tochihara, Y. (2014). A comparison of hydration effect on body fluid and temperature regulation between Malaysian and Japanese males exercising at mild dehydration in humid heat. *J. Physiol. Anthropol.* 33, 1–11. doi: 10.1186/1880-6805-33-5

Waldron, M., Villerius, V., and Murphy, A. (2015). Augmenting performance feedback does not affect 4 km cycling time-trials in the heat. *J. Sports Sci.* 33, 786–794. doi: 10.1080/02640414.2014.962579

Wang, F., Gao, C., Kuklane, K., and Holmer, I. (2013). Effects of various protective clothing and thermal environments on heat strain of unacclimated men: the PHS (predicted heat strain) model revisited. *Ind. Health* 51, 266–274. doi: 10.2486/indhealth.2012-0073

Watkins, S. L., Castle, P., Mauger, A. R., Sculthorpe, N., Fitch, N., Aldous, J., et al. (2014). The effect of different environmental conditions on the decision-making performance of soccer goal line officials. *Res. Sport. Med.* 22, 425–437. doi: 10.1080/15438627.2014.948624

Watson, P., Enever, S., Page, A., Stockwell, J., and Maughan, R. J. (2012). Tyrosine supplementation does not influence the capacity to perform prolonged exercise in a warm environment. *Int. J. Sport Nutr. Exerc. Metab.* 22, 363–373. doi: 10.1123/ijsnem.22.5.363

Webb, P. (1992). Temperatures of skin, subcutaneous tissue, muscle and core in resting men in cold, comfortable and hot conditions. *Eur. J. Appl. Physiol. Occup. Physiol.* 64, 471–476. doi: 10.1007/BF00625070

Webb, R. C., Bonifas, A. P., Behnaz, A., Zhang, Y., Yu, K. J., Cheng, H., et al. (2013). Ultrathin conformal devices for precise and continuous thermal characterization of human skin. *Nat. Mater.* 12, 938–944. doi: 10.1038/nmat3755

White, M. K., and Hodous, T. (1987). Reduced work tolerance associated with wearing protective clothing and respirators. *Am. Ind. Hyg. Assoc. J.* 48, 304–310. doi: 10.1080/15298668791384805

Wiggen, Ø. N., Waagaard, S. H., Heidelberg, C. T., and Oksa, J. (2013). Effect of cold conditions on double poling sprint performance of well-trained male cross-country skiers. *J. Strength Cond. Res.* 27, 3377–3383. doi: 10.1519/JSC.0b013e3182915e7d

Williamson, P. R., Lancaster, G. A., Craig, J. V., and Smyth, R. L. (2002). Meta-analysis of method comparison studies. *Stat. Med.* 21, 2013–2025. doi: 10.1002/sim.1158

Wright, H. E., McLellan, T. M., Larose, J., Hardcastle, S. G., Boulay, P., and Kenny, G. P. (2014). Are circulating cytokine responses to exercise in the heat augmented in older men? *Appl. Physiol. Nutr. Metab.* 39, 117–123. doi: 10.1139/apnm-2013-0223

Wright Beatty, H. E., Hardcastle, S. G., Boulay, P., Flouris, A. D., and Kenny, G. P. (2015). Increased air velocity reduces thermal and cardiovascular strain in young and older males during humid exertional heat stress. *J. Occup. Environ. Hyg.* 12, 625–634. doi: 10.1080/15459624.2015.1029613

Xu, X., Karis, A. J., Buller, M. J., and Santee, W. R. (2013). Relationship between core temperature, skin temperature, and heat flux during exercise in heat. *Eur. J. Appl. Physiol.* 113, 2381–2389. doi: 10.1007/s00421-013-2674-z

Yakovlev, V. V., and Utekhin, B. A. (1965). Errors in skin temperature measurements due to changes in evaporation under the sensor. *Bull. Exp. Biol. Med.* 60, 1210–1212.

Yakovlev, V. V., and Utekhin, B. A. (1966). Criteria for the evaluation of thermoelectric transducers and tests of certain designs of such transducers for measurements of skin temperature. *Gig. Sanit.* 31, 62–65.

Yeung, S. S., Ting, K. H., Hon, M., Fung, N. Y., Choi, M. M., Cheng, J. C., et al. (2016). Effects of cold water immersion on muscle oxygenation during repeated bouts of fatiguing exercise: a randomized controlled study. *Medicine (Baltimore)* 95:e2455. doi: 10.1097/MD.0000000000002455

Yogev, A., Hall, A. M., Jay, O., and White, M. D. (2015). Effects of elevated core temperature and normoxic 30% nitrous oxide on human ventilation during short duration, high intensity exercise. *Respir. Physiol. Neurobiol.* 206, 19–24. doi: 10.1016/j.resp.2014.10.005

Youhui, P., Zhiwei, L., and Li, P. (2010). Skin temperature measurement method. *J. Southeast Univ.* 26, 258–261.

Zhang, H., Arens, E., Huizenga, C., and Han, T. (2010). Thermal sensation and comfort models for non-uniform and transient environments: part I: local sensation of individual body parts. *Build. Environ.* 45, 380–388. doi: 10.1016/j.buildenv.2009.06.018

Zwolinska, M., and Bogdan, A. (2012). Impact of the medical clothing on the thermal stress of surgeons. *Appl. Ergon.* 43, 1096–1104. doi: 10.1016/j.apergo.2012.03.011

APPENDIX

The 172 articles included within objective 2, survey of use, are given below:

Adams et al., 2014, 2015; Almudehki et al., 2012; Altrichter et al., 2014; Amano et al., 2011; Ando et al., 2015; Bach et al., 2015a; Bain et al., 2011; Bain and Jay, 2011; Bandyopadhyay et al., 2011; Barr et al., 2011; Barwood et al., 2012, 2013, 2014, 2015, 2016; Best et al., 2012, 2014; Blacker et al., 2013; Bottoms and Price, 2014; Bourlai et al., 2012; Brade et al., 2014; Brown and Connolly, 2015; Bruning et al., 2013; Burdon et al., 2015; Byrne et al., 2011; Carlson et al., 2014; Caruso et al., 2015; Chou et al., 2011; Coombs et al., 2015; Corbett et al., 2015; Coso et al., 2011; Costello et al., 2015; Crabtree and Blannin, 2015; Cuddy et al., 2013, 2014, 2015; Daanen et al., 2015; Davey et al., 2013; De Pauw et al., 2011; De Sousa et al., 2014; Demachi et al., 2012a,b, 2013; Deren et al., 2012, 2014; Eijsvogels et al., 2014; Ely et al., 2011; Faulkner et al., 2015; Fernandes et al., 2014; Filho et al., 2015; Filingeri et al., 2014, 2015a,b; Flouris et al., 2015; Fonseca et al., 2015; Fortes et al., 2013; Gagnon et al., 2014a,b; Gagnon and Kenny, 2011; Gao et al., 2011, 2015, 2016; Gillis et al., 2015; Goh et al., 2011; Griggs et al., 2015; Guéritée et al., 2015a,b; Hayashi et al., 2012, 2014; Hillman et al., 2013; Hobson et al., 2013; Ihsan et al., 2013; Ito et al., 2013, 2015; James et al., 2014; Jovanović et al., 2014; Katica et al., 2011; Kenefick et al., 2011, 2014; Keramidas et al., 2013; Klute et al., 2014 Kounalakis et al., 2012; Lee et al., 2012, 2013, 2014a,b, 2015; Levels et al., 2012, 2013, 2014a,b; Limbaugh et al., 2013; Liu et al., 2013; MacRae et al., 2012, 2014; Matsuzuki et al., 2011; Mauger et al., 2014; Maughan et al., 2012; Mee et al., 2015; Mieras et al., 2014; Miyagawa et al., 2011; Moore et al., 2013; Mora-Rodriguez et al., 2011, 2012; Morris et al., 2014; Morrison et al., 2014a,b; Moyen et al., 2014a,b; Muller et al., 2011; Muñoz et al., 2012; Nassif et al., 2014; Niedermann et al., 2014; Noonan and Stachenfeld, 2012; Oda and Shirakawa, 2014; Oksa et al., 2014; Parker et al., 2013; Pearson et al., 2011; Périard et al., 2014; Pèriard et al., 2011, 2013; Périard and Racinais, 2015; Ping et al., 2011; Pryor et al., 2015; Pyke et al., 2015; Randall et al., 2015; Ravanelli et al., 2014; Richmond et al., 2015; Rintamäki et al., 2012; Roberge et al., 2012a,b; Sandsund et al., 2012; Schlader et al., 2011a,b,c; Shimazaki and Murata, 2015; Siegel et al., 2011; Sinclair and Brownsberger, 2013; Skein et al., 2012; Smith et al., 2013; Spitz et al., 2014; Stapleton et al., 2013; Sun et al., 2015; Suzuki et al., 2014; Temfemo et al., 2011; Teunissen et al., 2013, 2014; Tokizawa et al., 2014; Tsuji et al., 2012, 2014; Tumilty et al., 2011, 2014; Turner and Richards, 2015; Tyler and Sunderland, 2011a,b; Van Cutsem et al., 2015; Vanos et al., 2012; Veltmeijer et al., 2014; Wakabayashi et al., 2011, 2014; Waldron et al., 2015; Wang et al., 2013; Watkins et al., 2014; Watson et al., 2012; Wiggen et al., 2013; Wright et al., 2014; Wright Beatty et al., 2015; Xu et al., 2013; Yeung et al., 2016; Yogev et al., 2015; Zwolinska and Bogdan, 2012.