Thermographic imaging in sports and exercise medicine: a Delphi study and consensus statement on the measurement of human skin temperature

Danilo Gomes Moreira,[1, 25] Joseph T. Costello,[2] Ciro J. Brito,[3] Jakub G. Adamczyk,[4] Kurt Ammer,[5] Aaron J. E. Bach,[6] Carlos M. A. Costa,[7] Clare Eglin,[2] Alex A. Fernandes,[8] Ismael Fernández-Cuevas,[9, 25] José J. A. Ferreira,[10] Damiano Formenti,[11] Damien Fournet,[12] George Havenith,[13] Kevin Howell,[14] Anna Jung,[15] Glen P. Kenny,[16] Eleazar S. Kolosovas-Machuca,[17] Matthew J. Maley,[6] Arcangelo Merla,[18] David Pascoe,[19] Jose I. Priego-Quesada,[20] Robert G. Schwartz,[21] Adérito R. D. Seixas,[22] James Selfe,[23] Boris G. Vainer,[24] Manuel Sillero-Quintana[25]

[1] Federal Institute for Education, Science and Technology of Minas Gerais, Campus Governador Valadares, Brasil;

[2] Extreme Environments Laboratory, Department of Sport and Exercise Science, University of Portsmouth, Portsmouth, UK;

[3] Department of Physical Education, Federal University of Juiz de Fora, Governador Valadares, Brazil;

[4] Józef Piłsudski University of Physical Education in Warsaw, Warsaw, Poland;

[5] European Association of Thermology, Vienna, Austria; Medical Imaging Research Unit, University of South Wales, Pontypridd, United Kingdom;

[6] School of Exercise and Nutrition Sciences and Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, Australia;

[7] Aeronatutics Instruction and Adaptation Center, Minas Gerais, Brazil;

[8] Federal Institute for Education, Science and Technology of Minas Gerais, Campus Ipatinga, Brasil;

[9] School of Health and Sports Science, Cluster for Health Improvement, University of the Sunshine Coast, Australia;

[10] Department of Physical Therapy, Federal University of Paraíba, João Pessoa, Brazil;

[11] Department of Biomedical Sciences for Health, Università degli Studi di Milano, Italy;

[12] Thermal Sciences Laboratory, DECATHLON SportsLab, Villeneuve d'Ascq, France;

[13] Environmental Ergonomics Research Centre, Loughborough Design School, Loughborough University, Loughborough, United Kingdom;

[14] UCL Institute of Immunity and Transplantation, Royal Free Hospital, London, United Kingdom;

[15] Military Institute of Medicine, Warsaw, Poland;

[16] Human and Environmental Physiology Research Unit, School of Human Kinetics, University of Ottawa, Ottawa, Ontario, Canada;

[17] Autonomous University of San Luis Potosí, México;

[18] Department of Clinical Sciences and Bioimaging, School of Medicine, University of Chieti-Pescara, Chieti, Italy;

[19] Thermal Research Laboratory, Department of Kinesiology, Auburn University, Auburn, United States;

[20] Biophysics and Medical Physics Group, Department of Physiology, University of Valencia, Valencia, Spain;

[21] American Academy of Thermology. Piedmont Physical Medicine and Rehabilitation, PA, United States;

[22] Escola Superior de Saúde, Universidade Fernando Pessoa, Porto, Portugal;

[23] Manchester Metropolitan University, Manchester, United Kingdom;

[24] Rzhanov Institute of Semiconductor Physics SB RAS, Novosibirsk State University, Novosibirsk, Russia;

[25] Sports Department, Faculty of Sciences for Physical Activity and Sport (INEF), Technical University of Madrid, Madrid, Spain.

Corresponding author:

Danilo Gomes Moreira

Federal Institute for Education, Science and Technology of Minas Gerais, Campus Governador Valadares, Avenida Minas Gerais, nº 5189, Bairro Ouro Verde, Governador Valadares – MG, Brasil, CEP: 35057-760.

e-mail: danilo.moreira@ifmg.edu.br

Word count: 4091

Abstract

The importance of using infrared thermography (IRT) to assess skin temperature (t_{sk}) is increasing in clinical settings. Recently, its use has been increasing in sports and exercise medicine; however, no consensus guideline exists to address the methods for collecting data in such situations. The aim of this study was to develop a checklist for the collection of t_{sk} using IRT in sports and exercise medicine. We carried out a Delphi study to set a checklist based on consensus agreement from leading experts in the field. Panelists (n = 24) representing the areas of sport science (n=8; 33%), physiology (n=7; 29%), physiotherapy (n=3; 13%) and medicine (n=6; 25%), from 13 different countries completed the Delphi process. An initial list of 16 points was proposed which was rated and commented on by panelists in three rounds of anonymous surveys following a standard Delphi procedure. The panel reached consensus on 15 items which encompassed the participants' demographic information, camera/room or environment setup and recording/analysis of t_{sk} using IRT. The results of the Delphi produced the checklist entitled "Thermographic Imaging in Sports and Exercise Medicine (TISEM)" which is a proposal to standardize the collection and analysis of t_{sk} data using IRT. It is intended that the TISEM can also be applied to evaluate bias in thermographic studies and to guide practitioners in the use of this technique.

Keywords: Infrared thermography, guideline, protocol, checklist, thermoregulation.

Introduction

The growing importance of infrared thermography (IRT) measures of human skin temperature (t_{sk}) in health and disease has been evidenced by the increase in the number of publications with this technique. IRT is characterized by the use of a camera which can detect radiation and produce thermal images, called thermograms (Ring and Ammer, 2000). The thermograms contain temperature data that can be analyzed by specific software which provides temperature of a region of interest (ROI) (Costello et al., 2012b; Selfe et al., 2006). Common applications of IRT include: prevention and treatment of sports injuries (Hadžić et al., 2015; Hildebrandt et al., 2010), detection of delayed onset muscle soreness (Hani et al., 2012), evaluation of cryotherapy protocols (Adamczyk et al., 2016; Costello et al., 2012b; Selfe et al., 2014; Silva et al., 2017), assessment of brown adipose tissue activation (Robinson et al., 2016), evaluation of t_{sk} during cold (Fournet et al., 2013) and hot environment exposure (Gerrett et al., 2015; Vainer, 2005), following aerobic (Priego-Quesada et al., 2015b), anaerobic (Adamczyk et al., 2014) and resistance exercises (Ferreira et al., 2008). These applications have considerably growth in the use of IRT in recent years, due in part to improvements in the accuracy, functionality and affordability of camera technology, thus making IRT an emerging method of t_{sk} measurement in sports and exercise medicine (Bach et al., 2015b; Costello et al., 2012b; Fernandes et al., 2014).

IRT is a rapid emerging technique for the assessment of t_{sk} as it is versatile, non-invasive, wireless, and requires no contact with the individual (Bach et al., 2015b; Fernandes et al., 2014; Formenti et al., 2016). Due to its image capture capability, the selection of ROIs permits an evaluation of the t_{sk} distributions in different areas, consequently allowing its application in studies that require the analysis of several areas simultaneously (Fournet et al., 2013; Gerrett et al., 2015). Moreover, the thermograms allow the visualization of hot and cold areas. This has important implications in studies aimed at determining what location hot or cold t_{sk} is generated (Costello et al., 2012a; Eglin et al., 2013; Maley et al., 2014; Robinson et al., 2016; Selfe et al., 2010). Another advantage of IRT is the portability of cameras which can be used in a wide array of conditions and locations in both the laboratory and the field (Fernandes et al., 2016; Hildebrandt et al., 2012).

Previous reports have demonstrated poor agreement between t_{sk} measurements by IRT and contact devices (e.g. thermistors, thermocouples, iButtons) (Bach et al., 2015a; Bach et al., 2015b; Fernandes et al., 2014). However, there are some inherent limitations related to the methodology of IRT data acquisition that could act as confounding factors, thereby influencing temperatures outcomes (Fernández-Cuevas et al., 2015). For example, the distance the camera is from the subject and the room temperature of the laboratory where IRT recording is conducted can affect the data (Fernández-Cuevas et al., 2015). If thermograms from different distances (fields of view) of the same subject are compared, the variable number of pixels within the ROI can lead to inaccurate data (Ring and Ammer, 2000). Ammer (2015) compared the results of local thermograms with a total body thermogram and showed differences in the t_{sk} of the anterior thigh of up to 1.09 ± 0.93 °C (CI: 2.91; -0.74 °C). Likewise, a room without adequate temperature regulation can result in variable air temperatures thereby impacting results (Bach et al., 2015b). Most published studies using IRT have employed a temperature range of 18 °C to 25 °C (Fernández-Cuevas et al., 2015). However, it is well established that resting metabolic rate varies as a function of ambient temperature which can markedly influence thermoregulatory response (e.g. heat conversation or heat loss responses) and ultimate t_{sk} (Taylor et al., 2014). Consequently, it is important to ensure that ambient temperature conditions are adequately regulated to minimize any potential influence on the measurement of t_{sk} using IRT.

In order to prevent bias and improve the quality of data, several organizations have published their own protocols and quality control guidelines (Ammer, 2008; IACT, 2002; ISO, 2004; Mercer and Ring, 2009; Ring and Ammer, 2012; Ring and Ammer, 2000; Schwartz et al., 2006). Similarly, other investigators have discussed the need for standardization to ensure the quality of thermal image acquisition (Ammer, 2003; Ammer, 2015; Costello et al., 2012b; Ring et al., 2007a; Ring and Ammer, 2012). The acquisition of accurate t_{sk} data requires knowledge of the primary factors influencing the t_{sk} measurement. Fernández-Cuevas et al. (2015) defined the factors influencing the use of IRT in studies conducted on humans dividing them into environmental, technical and individual factors. However, no specific guidelines or checklist was provided. Given the wide array of factors that can affect the measurement of t_{sk} (Fernández-Cuevas et al., 2015), the use of IRT for scientific analysis can be challenging. However, by defining appropriate measurement standards and protocols, the accurate assessment of t_{sk} with IRT is possible. As reported by

Costello et al. (2012b), many researchers fail to report detailed information regarding the procedures and conditions under which IRT is employed. As a consequence, there can be a lack of standardization between studies which can affect the interpretability of the data. However, this limitation can be circumvented with the development of operational standards for the use of IRT.

In this context, we propose a checklist based on consensus agreement from leading experts in the field. Checklists tend to be more reliable than guidelines as they focus on the key points for simplicity and ease of application (Kelley et al., 2003). Similar checklists such as the PEDro (Verhagen et al., 1998) and CONSORT (Schulz et al., 2010) have been developed using expert consensus to improve methodological aspects of research and the contribution of these instruments is well established (Moher et al., 2001; Moseley et al., 2011). It is expected that a consensus instrument can contribute to the most appropriate application of IRT technique, including data collection and analysis. Therefore, this study aimed to develop a detailed checklist for the assessment of t_{sk} using IRT in sports and exercise medicine settings. It is intended to standardize the collection and analysis of t_{sk} data by end users which include clinicians, researchers and practitioners. This checklist could also be applied to evaluate bias in thermographic studies, and to guide practitioners in the use of this technique.

Methods

Participants

Of the 30 invited experts, 25 agreed to participate in the study. Two experts declined the invitation because they were too busy to participate, and a further three did not respond. One expert only completed the first round leaving a total of 24 experts who participated in the full evaluation. Panelists were selected based on their expertise in studies with IRT and thermal physiology, with at least three publications related to the discipline or area. Once a person was identified as an expert, an e-mail was sent to him/her with an invitation to participate; those who agreed received an electronic consent document. The panelists included experts who self-defined themselves as working predominantly in the sport sciences (n=8; 33%), physiology (n=7; 29%), physiotherapy (n=3; 13%) and medicine (n=6; 25%). They were currently working in academic and/or research institutions (n=17; 71%), practicing in hospitals as medical doctors (n=3; 13%), working for a company/industry (n=2; 8%), or working in the military (n=2; 8%). The panelists resided in the United Kingdom (n=5; 21%), Brazil (n=3; 13%), Australia (n=2; 8%), Italy (n=2; 8%), Poland (n=2; 8%), Spain(n=2; 8%), United States (n=2; 8%), Austria (n=1; 4%), Canada (n=1; 4%), France (n=1; 4%), Mexico (n=1; 4%), Portugal (n=1; 4%) and Russia (n=1; 4%). Panelists had a wealth of experience working with IRT, thermoregulation, and the assessment of t_{sk} (Median=8 years; range 4-32), and had published a median of eight (range 3-80) full peer reviewed articles related to the subject examined in the current manuscript. In addition, a search on the Scopus database on 06/26/2017 showed an average H index of 9 (range: 3 to 35). Participants received an information document describing the study with consent indicated by completion of the Delphi survey. The first, second, third and last authors organized the work of the panel as part of the core research team but did not participate as panel members. The core panel was responsible for the development of the initial items, the analysis, organization and reporting of the decisions, as well as communicating with the panelists.

Research Design

A Delphi procedure was applied in the present study, as previously described (Boulkedid et al., 2011; Dalkey and Helmer, 1963; Hsu and Sandford, 2007; Steurer, 2011). The Delphi procedure is based on developing a consensus among a group of experts through a series of questionnaires interspersed with controlled feedback (Whiting et al., 2003). In this procedure the expert evaluation, judgment, phrasing and scoring of each round is completed independently. As previously demonstrated, controlled and anonymous feedback also helps the experts to gain a consensus (Boulkedid et al., 2011; Dalkey and Helmer, 1963; Hsu and Sandford, 2007; Steurer, 2011). To initiate the process, a literature review was conducted in March 2016 to identify the available guidelines. Subsequently, instead of asking open questions to the panelists, an initial list of items for inclusion in the checklist was developed. This approach is considered appropriate if basic information is already available within the literature (Hsu and Sandford, 2007). In addition, as infrared technology has been improved in recent years, only documents published since 2000 were included (for further details and a full rationale see Costello et al. (2013) and Bach et al. (2015b)). Therefore, nine documents (Ammer, 2008; Fernández-Cuevas et al., 2015; IACT, 2002; ISO, 2004; Mercer and Ring, 2009; Ring and Ammer, 2000; Ring et al., 2007b; Schwartz et al., 2006; Schwartz et al., 2015) which contained recommendations on the use of IRT for measurements in humans were used to develop an itemized list.

The core panel (authors DGM, JTC, CJB and MSQ) reviewed the nine documents to summarize the empirical and theoretical evidence regarding the procedures involving IRT. A preliminary list (with 16 items for inclusion in the checklist) was settled based on scientific evidence provided in the selected documents. The core panel previously established that only aspects relating to the measurement of t_{sk} using IRT would be included in the checklist, and other aspects in the applications of IRT were not included.

Although no criterion is universally accepted to address consensus, the value of 80% agreement is mostly used (Bahl et al., 2016; Boulkedid et al., 2011; Hsu and Sandford, 2007; Snyder et al., 2014; Whiting et al., 2003). A five point Likert Scale (strongly agree, moderately agree, neutral, moderately disagree, strongly disagree) (Chipchase et al., 2012; Whiting et al., 2003) was used in all rounds to rate each item for inclusion in the checklist. Therefore, we applied the criteria of 80% of the sum of responses of ‘strongly agree’ and ‘moderately agree’ to approve an item. To help panel members in their decision-making, the core panel summarized the evidence and provided the references which supported each item. In addition, the panelists were encouraged to identify any study or practical experience that could help in the discussion. In every round, all members were given an opportunity to comment on the items and suggest possible rephrasing. The panelists had 15 days to respond to each round and all communication was conducted by electronic mail.

The responses of each Delphi round were organized to include the results of the previous round and a summary of all panel members’ evaluation. All modifications were shared in the subsequent rounds as previously reported (Chipchase et al., 2012; Snyder et al., 2014; Whiting et al., 2003). Furthermore, the decision made in each round comprised of six distinct actions: (1) *Modify*: when an item was substantially modified, either by suggestion of a panel member or to meet new evidence; (2) *Rephrase*: when an item was rephrased to improve understanding without changing the meaning; (3) *Divide*: when an item was divided into two different items; (4) *Exclude*: when an item did not meet the criteria and was excluded from the checklist; (5) *Include*: when a panel member suggested a new item; and (6) *Approve*: when an item met the criteria and was approved to be part of the checklist. Consensus was considered to be achieved for an item if: a) the criteria of 80% of the sum of responses of ‘strongly agree’ and ‘moderately agree’ was reached; and b) no panelist recommended a relevant change in wording or provided new or additional evidence. All comments regarding rephrasing were incorporated when revising the checklist.

Round 1

The initial list of possible items for inclusion was sent to all panel members which included information on the the aim of the consensus protocol and the process. The scientific reasoning for the inclusion of each item was presented. All panelists were asked to analyze each item while considering two main points: (1) validity, that was defined as the extent to which the characteristics of the item are appropriate to the objective of the checklist (Boulkedid et al., 2011); and (2) feasibility, which was defined as the practical viability of the item (Boulkedid et al., 2011). The panelists were encouraged to support their answers with information based on peer-reviewed literature whenever possible. When this was not possible, it was suggested that they justify their positions using their practical experience in the use of thermography.

Round 2

The results of round one were analyzed by the core panel and a report was prepared containing the response rating, as well as a summary of all the comments received for each item. In addition, in this round, panelists were asked to indicate which tense they preferred the checklist to be in, and whether they wanted to propose a new item not included in the initial list. In this context, the following questions were asked: 1) In order to improve the understanding and interpretation of the proposed items, which tense do you deem to be the most appropriate for the presentation of the checklist? 2) Do you want to propose an item that is not included in this checklist? If yes, please, propose the item and identify where you think it should be located in the checklist.

Round 3

Based on the results of round two, all of the items were modified according to the chosen tense (have/should/must). As some items were approved in round two (5 items), only those items that did not receive consensus were assessed further. The response ratings and a summary of all the comments received were presented again. The study was concluded in this round since we achieved the previously established threshold for consensus. The phases of the study are presented in Figure 1.

*** Insert Figure 1 here***

Results

Round 1

Decisions in this round included modification (6; 38%), rephrasing (8; 50%), and exclusion (2; 12%) of items. In general, the items obtained a high level of agreement between the experts, however some items were judged incomplete or wrong (the number of the approved item in the final checklist is expressed in table 1 – items: individual data, previous instructions, environmental condition, image background, acclimation and camera preparation) while others required alternate phrasing (items: extrinsic factors, environmental setup, equipment, image recording, camera position, emissivity, body position and image evaluation). For the incomplete items, new evidence was provided by panel members and subsequently incorporated. Likewise, suggested grammatical edits were incorporated to improve clarity. Although some items met the approval criteria, none was approved in this round, since it was deemed by the panelists that the proposed edits should be re-evaluated.

The core panel indentified two items (assessment time and method of drying the skin) to be removed since they were judged not relevant to the checklist. Five panelists argued that both items were not related to the objective of the checklist and therefore they should be excluded. All edits were highlighted and explained in subsequent round for evaluation.

Round 2

The decisions in this round comprised modification (4; 25%), rephrasing (6; 38%), division (1; 6%) and approval (5; 31%). Five items were approved since they met the approval criteria and no further relevant information was provided by the panelists (items: extrinsic factor, environmental condition, camera position, emissivity and image evaluation). Ten panelists suggested a new approach to address the items “assessment time” and “method of drying the skin”, which were suggested to be removed in the previous round because they were not related to the goal of the checklist. Based on the feedback received on these items, a new version was proposed to make them applicable. In addition, the item “assessment time” was divided into “assessment time” and “assessment operators”, since most of the comments indicated that two distinct aspects were addressed. Regarding the question about which tense would be the most appropriate, the majority of panel members selected should/has/must sentences (15;63%) followed by past tense (4;17%), question sentences (4;17%), and present tense (1;4%). On the basis of these data, the items on the checklist were modified.

Only one panel member proposed a new item regarding aspects that should be considered when presenting IRT images in scientific articles. The core panel considered this item related to the category “body position”, and it was therefore added to item 13. All decisions were communicated and submitted to the next round.

Round 3

The items approved in the previous round were not included in this round. Thus, the decisions involved approval (10; 92%) and exclusion (2; 8%). Two items (image background and assessment operators) did not meet the criteria for inclusion and were subsequently excluded from the checklist. The responses from each of the rounds, and the decisions made by the core panel are shown in table 1.

Insert table 1 here

Final document

The final Checklist was structured as a list of 15 items (table 2) which should be marked “yes”, “no”, or “unclear”.

Insert Table 2 here

Discussion

In the absence of a consensus guideline to measure t_{sk} with IRT in sports and exercise medicine settings, a Delphi procedure was applied to develop a checklist for addressing the methodological aspects of data collection. The final checklist entitled “Thermographic Imaging in Sports and Exercise Medicine” (TISEM) contains 15 items which were approved by 24 world leading experts covering different fields of expertise (e.g. sport sciences, physiology, physiotherapy and medicine). The items encompassed the participants’ demographic information (items 1, 2 and 3), camera/room or environment setup (items 4, 5, 6, 7, 8, 9, 10 and 11) and recording/analysis (items 12, 13, 14 and 15) of t_{sk} using IRT. Considering the rapidly emerging use of IRT in sports and exercise medicine (Costello et al., 2013; Hildebrandt et al., 2010; Hildebrandt et al., 2012), TISEM addresses the relevant issues regarding the methodological aspects of data collection. The items from TISEM were organized to identify key methodological aspects regarding the use of infrared thermography in humans, especially when conducting research investigations. Similarly, it is proposed that this checklist could be used to collect t_{sk} data in a medical or clinical setting, since all items are equally relevant in non-sports and exercise medicine setting.

It is well established that age (Kenny and Journeay, 2010), sex (Gagnon and Kenny, 2012), body composition (Chudecka et al., 2015; Savastano et al., 2009), ethnicity (Maley et al., 2014), prevalence of smoking (Bornmyr and Svensson, 1991; Ijzerman et al., 2003), and others impact thermoregulation and t_{sk}. As such, we advise physical characteristics (e.g. age, sex, body mass, height and body mass index), as well as ethnicity and smoking history be reported. A number of studies have previously demonstrated the effect of physical fitness on t_{sk} (Abate et al., 2013; Akimov and Son'kin, 2011; Chudecka and Lubkowska, 2010; Formenti et al., 2013; Quesada et al., 2015a), therefore it is recommended that participants' physical activity profile (e.g. frequency, duration, intensity, and activity description) and/or physical fitness (e.g. aerobic capacity) be reported. Moreover, although the exact time frames of the impact of alcoholic beverages, smoking, caffeine, large meals, ointments, cosmetics, showering and sunbathing may have on t_{sk} are not well known (Fernández-Cuevas et al., 2015; Ring and Ammer, 2000), TISEM advises that users control these variables in order to standardize the data collection procedures within and, where possible, between studies. These variables should be confirmed verbally before the assessment and the use of any medicinal treatments or drugs should be recorded. Additionally, t_{sk} is heavily influenced by extrinsic factors such as prior physical activity (Bach et al., 2015a; Formenti et al., 2016; Tanda, 2016), as well as physical or medical treatments such as massage (Adamczyk et al., 2016; IACT, 2002; Ring and Ammer, 2000), electrotherapy (Ring and Ammer, 2000), ultrasound (Fernández-Cuevas et al., 2015), and by heat (Bach et al., 2015a; Gerrett et al., 2014) or cold (Chesterton et al., 2002; Vellard and Arfaoui, 2016) exposure. Therefore, any intervention prior to or during the assessment of t_{sk} should be recorded and detailed.

In relation to the experimental setup and reporting of environmental conditions it is important to describe the ambient temperature and relative humidity where the assessment took place. Skin temperature is influenced by the environment, especially if the skin is exposed (Bach et al., 2015b; Fernandes et al., 2014), therefore TISEM recommends that mean ambient temperature (°C; ± standard deviation) and relative humidity (%; ± standard deviation) be reported. Similarly, external factors, such as infrared radiation (e.g. electronic devices, lightning) or airflow (e.g. under an air conditioning unit, ceiling fans, open windows/doors), are likely to impact on and interact with t_{sk}. Therefore, TISEM suggests completing the assessment away from these factors and reporting whether exposure to any of these conditions was unavoidable. Since a large number of infrared camera models are currently available including cooled and uncooled cameras (Bach et al., 2015b; Ring and Ammer, 2000), TISEM advises that manufacturer, model and accuracy of the camera is detailed. Where available, it is also important to report when and where the camera was last calibrated. Regarding the sensor stabilization of IRT cameras, depending on the technology, some models need to be turned on for some time prior to the assessment in order to ensure consistent readings. To determine the time frame to address this issue, TISEM recommends following manufacturer's guidelines or performing a quality assurance test, as described by Ring et al. (2007b). In addition, when baseline measurements of t_{sk} are required, TISEM recommends that an acclimation period be conducted in the examination room wherein ambient temperature and humidity are regulated. Although previous research (Marins et al., 2014) showed that a 10-min acclimation period is sufficient wherein differences between external and internal (testing room) temperature is less than 5 ºC, the panelists agreed (63%) that 15 minutes should be the minimum recommended acclimation period (Fernández-Cuevas et al., 2015; IACT, 2002; Ring and Ammer, 2000). However, it is important to note that extreme temperature (e.g. more than 20 ºC of difference between external temperature and room temperature), can require more time to acclimatize because of the marked effects on t_{sk} (Taylor et al., 2014). Moreover, the after-effect of the removal of clothing should be considered given that it can influence t_{sk} up to 20 minutes after undressing (Vainer, 2001). Therefore, the time used must be determined in accordance with the objective of the study/assessment, the ROI (open skin or that under clothes) and the environmental conditions. The checklist also advises that the distance between the skin and camera and percentage of the region of interest within the image should be detailed in order to guarantee reproducibility across studies. As demonstrated by Ammer (2015), when the camera is placed close to the region of interest, the field of view provides a more detailed temperature information. In addition, the camera should be positioned perpendicular to the region of interest, otherwise the assessment can result in a critical loss of information (Tkacova et al., 2010).

While controversy in the literature regarding emissivity exists (Sanchez-Marin et al., 2009; Steketee, 1973), 96% of experts strongly agreed that an emissivity of 0.98 (ε) should be used for clean dry skin. Due to circadian rhythm (Costa et al., 2015; Marins et al., 2015), t_{sk} is likely to change during the course of the day, therefore, when individuals are assessed over multiple days or when comparisons between participants are made at different time of the day, the time of day at which the images were recorded should be reported. The panel of experts agreed that the use of a standardized body position as well as the selection of regions of interest should be sufficiently described to ensure the reliability and reproducibility of the data. Moreover, the presentation of a visual example is recommended as it may add important information or representation about how regions of interests were defined. Since water impacts on emissivity (Fernández-Cuevas et al., 2015) in some situations, particularly during cold water immersion or exercising in water, the experts agreed (92%) that the skin could be dried. However, the method of drying (e.g. towel patting) should be clearly described and reported. Recent recommendations by Seixas et al. (2014) suggested that the skin should be carefully dried with a microfiber towel to limit irritation of the skin (that may occur with more abrasive fabrics). At the same time, it is quantitatively demonstrated that moisturizing the skin affects t_{sk} contrast and may be used to enhance the surface vessels thermal pattern (Vainer, 2001). In addition, a suitable practice within extremity cooling studies is to use a thin plastic bag to prevent the extremity from becoming wet (Eglin et al., 2013; Maley et al., 2014). Finally, the method of analysis including the software used and whether or not the analysis was completed manually or automatically should be described. Similarly, the method employed to calculate the final temperature value (e.g. average, median, maximum or minimum) should be clear described. As much information about the process itself should be provided so that others can replicate the findings if needed.

The number of experts who have completed the process (24) is greater than other studies using a similar methodological design (Boulkedid et al., 2011). In addition the experience of the panel experts (median = 8 years; range 3 to 32 years and publications median = 8; range 3 to 80), the multiple nationalities (13) and different professional backgrounds (4) of the panelists allow a thorough and broad analysis of the use of IRT, illustrating the positive characteristics of the current Delphi procedure (Boulkedid et al., 2011; Hsu and Sandford, 2007). However, the study is not without limitations. While the Delphi panel consisted of experts representing a range of disciplines, the process would have benefitted from a greater inclusion of practitioners that use IRT daily. In addition, the invitation of panelists could have caused selection bias. Because the recommendations are primarily based on experts' opinion, users should take into account the possibility of the checklist does not address all issues. In this sense, additional items may be required since scientific evidence has become available.

Conclusion

We have provided a checklist with 15 items directed at standardizing the assessment of t_{sk} using IRT for a wide array of end-users including practitioners, sports scientists, exercise physicians, medical professionals and others. This checklist is not limited to this setting, and may also be used in others fields such occupational medicine and public health. It is intended that the TISEM can also be applied to evaluate bias in thermographic studies, and to guide practitioners in the use of this technique.

Acknowledgements

National Council of Scientific and Technologic Development (CNPq) for the phD Scholarship number 205815/2014-6 for DGM and 234243/2014-7 for CJB. Participating societies included the European European Association of Thermology (EAT), American Academy of Thermology (AAT), and Polish Society of Medical Thermography.

Competing interests

None declared.

References

Abate, M., Di Carlo, L., Di Donato, L., Romani, G., Merla, A., 2013. Comparison of cutaneous termic response to a standardised warm up in trained and untrained individuals. J Sports Med Phys Fitness 53, 209-215.

Adamczyk, J.G., Boguszewski, D., Siewierski, M., 2014. Thermographic evaluation of lactate level in capillary blood during post-exercise recovery. Kinesiology 46, 186-193.

Adamczyk, J.G., Krasowska, I., Boguszewski, D., Reaburn, P., 2016. The use of thermal imaging to assess the effectiveness of ice massage and cold-water immersion as methods for supporting post-exercise recovery. J Therm Biol 60, 20-25.

Akimov, E., Son'kin, V., 2011. Skin temperature and lactate threshold during muscle work in athletes. Hum Physiol 37, 621-628.

Ammer, K., 2003. Need for standardisation of measurements in thermal imaging. Thermography and Lasers in Medicine. Akademickie Centrum Graficzno-Marketigowe Lodart SA, Lodz, 13-18.

Ammer, K., 2008. The Glamorgan Protocol for recording and evaluation of thermal images of the human body. Thermol Int 18, 125-129.

Ammer, K., 2015. Do we need reference data of local skin temperatures? Thermol Int 25, 45-47.

Bach, A.J., Stewart, I.B., Disher, A.E., 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.

Bach, A.J.E., Stewart, I.B., Minett, G.M., Costello, J.T., 2015b. Does the technique employed for skin temperature assessment alter outcomes? A Systematic Review. Physiol Meas 36 (9), 27-51.

Bahl, J.S., Dollman, J., Davison, K., 2016. The development of a subjective assessment framework for individuals presenting for clinical exercise services: A Delphi study. J Sci Med Sport 19 (11), 872-876.

Bornmyr, S., Svensson, H., 1991. Thermography and laser- Doppler flowmetry for monitoring changes in finger skin blood flow upon cigarette smoking. Clin Physiol 11 (2), 135-141.

Boulkedid, R., Abdoul, H., Loustau, M., Sibony, O., Alberti, C., 2011. Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. PLoS One 6, e20476.

Chesterton, L.S., Foster, N.E., Ross, L., 2002. Skin temperature response to cryotherapy. Phys Med Rehabil 83 (4), 543-549.

Chipchase, L., Schabrun, S., Cohen, L., Hodges, P., Ridding, M., Rothwell, J., Taylor, J., Ziemann, U., 2012. A checklist for assessing the methodological quality of studies using transcranial magnetic stimulation to study the motor system: an international consensus study. Clin Neurophysiol 123 (9), 1698-1704.

Chudecka, M., Lubkowska, A., 2010. Temperature changes of selected body's surfaces of handball players in the course of training estimated by thermovision, and the study of the impact of physiological and morphological factors on the skin temperature. J Therm Biol 35, 379-385.

Chudecka, M., Lubkowska, A., Leźnicka, K., Krupecki, K., 2015. The Use of Thermal Imaging in the Evaluation of the Symmetry of Muscle Activity in Various Types of Exercises (Symmetrical and Asymmetrical). J Hum Kinet 49 (11), 141-147.

Costa, C.M., Sillero-Quintana, M., Pinonosa Cano, S., Moreira, D.G., Brito, C.J., Fernandes, A.A., Pussieldi, G.A., Marins, J.C., 2015. Daily oscillations of skin temperature in military personnel using thermography. J R Army Med Corps 162 (5), 335-342.

Costello, J., Stewart, I.B., Selfe, J., Karki, A.I., Donnelly, A.E., 2013. Use of thermal imaging in sports medicine research: a short report: short article. Int SportMed Journal 14 (2), 94-98.

Costello, J.T., Culligan, K., Selfe, J., Donnelly, A.E., 2012a. Muscle, skin and core temperature after −110 ºC cold air and 8 ºC water treatment. PloS One 7, e48190.

Costello, J.T., McInerney, C.D., Bleakley, C.M., Selfe, J., Donnelly, A.E., 2012b. The use of thermal imaging in assessing skin temperature following cryotherapy: a review. J Therm Biol 37, 103-110.

Dalkey, N., Helmer, O., 1963. An experimental application of the Delphi method to the use of experts. Manag Sci 9 (3), 458-467.

Eglin, C.M., Golden, F.S., Tipton, M.J., 2013. Cold sensitivity test for individuals with non-freezing cold injury: the effect of prior exercise. Extrem Physiol Med 2 (1), 1-8.

Fernandes, A.A., Amorim, P.R.S., Brito, C.J., Moura, A.G., Moreira, D.G., Costa, C.M.A., Sillero-Quintana, M., Marins, J.C.B., 2014. Measuring skin temperature before, during and after exercise: a comparison of thermocouples and infrared thermography. Physiol Meas 35 (2), 189-203.

Fernandes, A.A., Moreira, D.G., Brito, C.J., da Silva, C.D., Sillero-Quintana, M., Pimenta, E.M., Bach, A.J., Garcia, E.S., Marins, J.C.B., 2016. Validity of inner canthus temperature recorded by infrared thermography as a non-invasive surrogate measure for core temperature at rest, during exercise and recovery. J Therm Biol 62, 50-55.

Fernández-Cuevas, I., Marins, J.C.B., Lastras, J.A., Carmona, P.M.G., Cano, S.P., García-Concepción, M.Á., Sillero-Quintana, M., 2015. Classification of factors influencing the use of infrared thermography in humans: a review. Infrared Phys Techn 71, 28-55.

Ferreira, J.J., Mendonca, L.C., Nunes, L.A., Andrade Filho, A.C., Rebelatto, J.R., Salvini, T.F., 2008. Exercise-associated thermographic changes in young and elderly subjects. Ann Biomed Eng 36, 1420-1427.

Formenti, D., Ludwig, N., Gargano, M., Gondola, M., Dellerma, N., Caumo, A., Alberti, G., 2013. Thermal imaging of exercise-associated skin temperature changes in trained and untrained female subjects. Ann Biomed Eng 41, 863-871.

Formenti, D., Ludwig, N., Trecroci, A., Gargano, M., Michielon, G., Caumo, A., Alberti, G., 2016. Dynamics of thermographic skin temperature response during squat exercise at two different speeds. J Therm Biol 59, 58-63.

Fournet, D., Ross, L., Voelcker, T., Redortier, B., Havenith, G., 2013. Body mapping of thermoregulatory and perceptual responses of males and females running in the cold. J Therm Biol 38, 339-344.

Gagnon, D., Kenny, G.P., 2012. Does sex have an independent effect on thermoeffector responses during exercise in the heat? J Physiol 590, 5963-5973.

Gerrett, N., Ouzzahra, Y., Coleby, S., Hobbs, S., Redortier, B., Voelcker, T., Havenith, G., 2014. Thermal sensitivity to warmth during rest and exercise: a sex comparison. Eur J Appl Physiol 114, 1451-1462.

Gerrett, N., Ouzzahra, Y., Redortier, B., Voelcker, T., Havenith, G., 2015. Female thermal sensitivity to hot and cold during rest and exercise. Physiol. Behav 152, 11-19.

Hadžić, V., Širok, B., Malnəršič, A., Čoh, M., 2015. Can infrared thermography be used to monitor fatigue during exercise? A case study. J Sport Health 1-4.

Hani, H.A.-N., Jerrold, S.P., Michael, S.L., Lee, S.B., 2012. The use of thermal infra-red imaging to detect delayed onset muscle soreness. J Vis Exp 59, e3551.

Hildebrandt, C., Raschner, C., Ammer, K., 2010. An overview of recent application of medical infrared thermography in sports medicine in Austria. Sensors 10, 4700-4715.

Hildebrandt, C., Zeilberger, K., Ring, E.F.J., Raschner, C., 2012. The application of medical Infrared Thermography in sports medicine, in: Zaslav, K.R. (Ed.), An International Perspective on Topics in Sports Medicine and Sports Injury. InTech, p. 534.

Hsu, C.-C., Sandford, B.A., 2007. The Delphi technique: making sense of consensus. Practical Assessment, Research & Evaluation 12, 1-8.

IACT, 2002. Thermology Guidelines. Standards and protocolos in Clinical Thermography Imaging. http://www.iact-org.org/professionals/thermog-guidelines.html. Acessed 01 Feb 2015.

Ijzerman, R.G., Serne, E.H., van Weissenbruch, M.M., de Jongh, R.T., Stehouwer, C.D., 2003. Cigarette smoking is associated with an acute impairment of microvascular function in humans. Clin Sci 104, 247-252.

ISO, 2004. Ergonomics-Evaluation of thermal strain by physiological measurements. International Standardization Organization Geneva.

Kelley, K., Clark, B., Brown, V., Sitzia, J., 2003. Good practice in the conduct and reporting of survey research. Int J Qual Health Care 15, 261-266.

Kenny, G.P., Journeay, W.S., 2010. Human thermoregulation: separating thermal and nonthermal effects on heat loss. Front Biosci 15, 259-290.

Maley, M.J., Eglin, C.M., House, J.R., Tipton, M.J., 2014. The effect of ethnicity on the vascular responses to cold exposure of the extremities. Eur J Appl Physiol 114, 2369-2379.

Marins, J.C.B., Formenti, D., Costa, C.M.A., Fernandes, A.d.A., Sillero-Quintana, M., 2015. Circadian and gender differences in skin temperature in militaries by thermography. Infrared Phys Techn 71, 322-328.

Marins, J.C.B., Moreira, D.G., Cano, S.P., Sillero-Quintana, M., Soares, D.D., Fernandes, A.A., Silva, F.S., Costa, C.M.A., Amorim, P.R.S., 2014. Time required to stabilize thermographic images at rest. Infrared Phys Technol 65, 30-35.

Mercer, J.B., Ring, E.F.J., 2009. Fever screening and infrared thermal imaging: concerns and guidelines. Thermol Int 19, 67-69.

Moher, D., Jones, A., Lepage, L., Group, C., 2001. Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. JAMA 285, 1992-1995.

Moseley, A.M., Herbert, R.D., Maher, C.G., Sherrington, C., Elkins, M.R., 2011. Reported quality of randomized controlled trials of physiotherapy interventions has improved over time. J Clin Epidemiol 64, 594-601.

Priego-Quesada, J.I., Carpes, F.P., Bini, R.R., Palmer, R.S., Pérez-Soriano, P., de Anda, R.M.C.O., 2015a. Relationship between skin temperature and muscle activation during incremental cycle exercise. J Therm Biol 48, 28-35.

Priego-Quesada, J.I., Guillamón, N.M., de Anda, R.M.C.O., Psikuta, A., Annaheim, S., Rossi, R.M., Salvador, J.M.C., Pérez-Soriano, P., Palmer, R.S., 2015b. Effect of perspiration on skin temperature measurements by infrared thermography and contact thermometry during aerobic cycling. Infrared Phys Technol 72, 68-76.

Ring, E., Ammer, K., Wiecek, B., Plassmann, P., Jones, C., Jung, A., Murawski, P., 2007a. Quality assurance for thermal imaging systems in medicine. Thermol Int 17, 103-106.

Ring, E.F., Ammer, K., 2012. Infrared thermal imaging in medicine. Physiol Meas 33, R33-46.

Ring, E.F.J., Ammer, K., 2000. The Technique of Infra red Imaging in Medicine. Thermol Int 10, 7-14.

Ring, F.J., Ammer, K., Wiecek, B., Plassmann, P., Jones, C.D., Jung, A., Murawski, P., 2007b. Quality assurance of thermal imaging systems in medicine. Thermol Int 16, 10-15.

Robinson, L.J., Law, J.M., Symonds, M.E., Budge, H., 2016. Brown adipose tissue activation as measured by infrared thermography by mild anticipatory psychological stress in lean healthy females. Exp Physiol 101, 549–557.

Sanchez-Marin, F.J., Calixto-Carrera, S., Villaseñor-Mora, C., 2009. Novel approach to assess the emissivity of the human skin. J Biomed Opt 14, 024006.

Savastano, D.M., Gorbach, A.M., Eden, H.S., Brady, S.M., Reynolds, J.C., Yanovski, J.A., 2009. Adiposity and human regional body temperature. Am J Clin Nutr 90, 1124-1131.

Schulz, K.F., Altman, D.G., Moher, D., 2010. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMC medicine 340, c332.

Schwartz, R., Elliott, R., Goldberg, G., 2006. Guidelines for neuromusculoskeletal thermography. Thermol Int 16, 5-9.

Schwartz, R.G., O'Young, B., Getson, P., Govindan, S., Uricchio, J., Bernton, T., Brioschi, M., Zhang, H.-Y., 2015. Guidelines for neuromusculoskeletal infrared thermography sympathetic skin response (ssr) studies. Pan Am J Med Thermol 2, 35-43.

Seixas, A., Gonjo, T., Vardasca, R., Gabriel, J., Fernandes, R., Vilas-Boas, J., 2014. A preliminary study on the relationship between energy expenditure and skin temperature in swimming, 12th International Conference on Quantitative InfraRed Thermography, Bordeaux, France, pp. 90-97.

Selfe, J., Alexander, J., Costello, J.T., May, K., Garratt, N., Atkins, S., Dillon, S., Hurst, H., Davison, M., Przybyla, D., 2014. The effect of three different (-135 C) whole body cryotherapy exposure durations on elite rugby league players. PLoS One 9, e86420.

Selfe, J., Hardaker, N., Thewlis, D., Karki, A., 2006. An accurate and reliable method of thermal data analysis in thermal imaging of the anterior knee for use in cryotherapy research. Arch Phys Med Rehabil 87, 1630-1635.

Selfe, J., Sutton, C., Hardaker, N., Greenhalgh, S., Karki, A., Dey, P., 2010. Cold Females, A Distinct Group Of Patellofemoral Pain Syndrome Patients? J Orthop Sports Physical 40, A42.

Silva, Y.A., Santos, B.H., Andrade, P.R., Santos, H.H., Moreira, D.G., Sillero-Quintana, M., Ferreira, J.J.A., 2017. Skin temperature changes after exercise and cold water immersion. Sport Sci Health 13, 195-202.

Snyder, K.R., Evans, T.A., Neibert, P.J., 2014. Developing a Framework for Ankle Function: A Delphi Study. J Athl Train 49, 747-757.

Steketee, J., 1973. Spectral emissivity of skin and pericardium. Phys Med Biol 18, 686-694.

Steurer, J., 2011. The Delphi method: an efficient procedure to generate knowledge. Skeletal Radiology 40, 959-961.

Tanda, G., 2016. Skin temperature measurements by infrared thermography during running exercise. Exp Therm Fluid Sci 71, 103-113.

Taylor, N.A., Tipton, M.J., Kenny, G.P., 2014. Considerations for the measurement of core, skin and mean body temperatures. J Therm Biol 46, 72-101.

Tkacova, M., Hudak, R., Foffova, P., Zivcak, J., 2010. An importance of camera - subject distance and angle in musculoskeletal application of medical thermography. Acta Electrotechnica et Informatica 10, 57-60.

Vainer, B.G., 2001. Treated-skin temperature regularities revealed by IR thermography, Aerospace/Defense Sensing, Simulation, and Controls. International Society for Optics and Photonics, pp. 470-481.

Vainer, B.G., 2005. FPA-based infrared thermography as applied to the study of cutaneous perspiration and stimulated vascular response in humans. Phys Med Biol 50, R63-94.

Vellard, M., Arfaoui, A., 2016. Detection by Infrared Thermography of the Effect of Local Cryotherapy Exposure on Thermal Spreadin Skin. J Imaging 2 (2), 1-9.

Verhagen, A.P., de Vet, H.C., de Bie, R.A., Kessels, A.G., Boers, M., Bouter, L.M., Knipschild, P.G., 1998. The Delphi list: a criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. J Clin Epidemiol 51, 1235-1241.

Whiting, P., Rutjes, A.W., Reitsma, J.B., Bossuyt, P.M., Kleijnen, J., 2003. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol 3, 3-25.

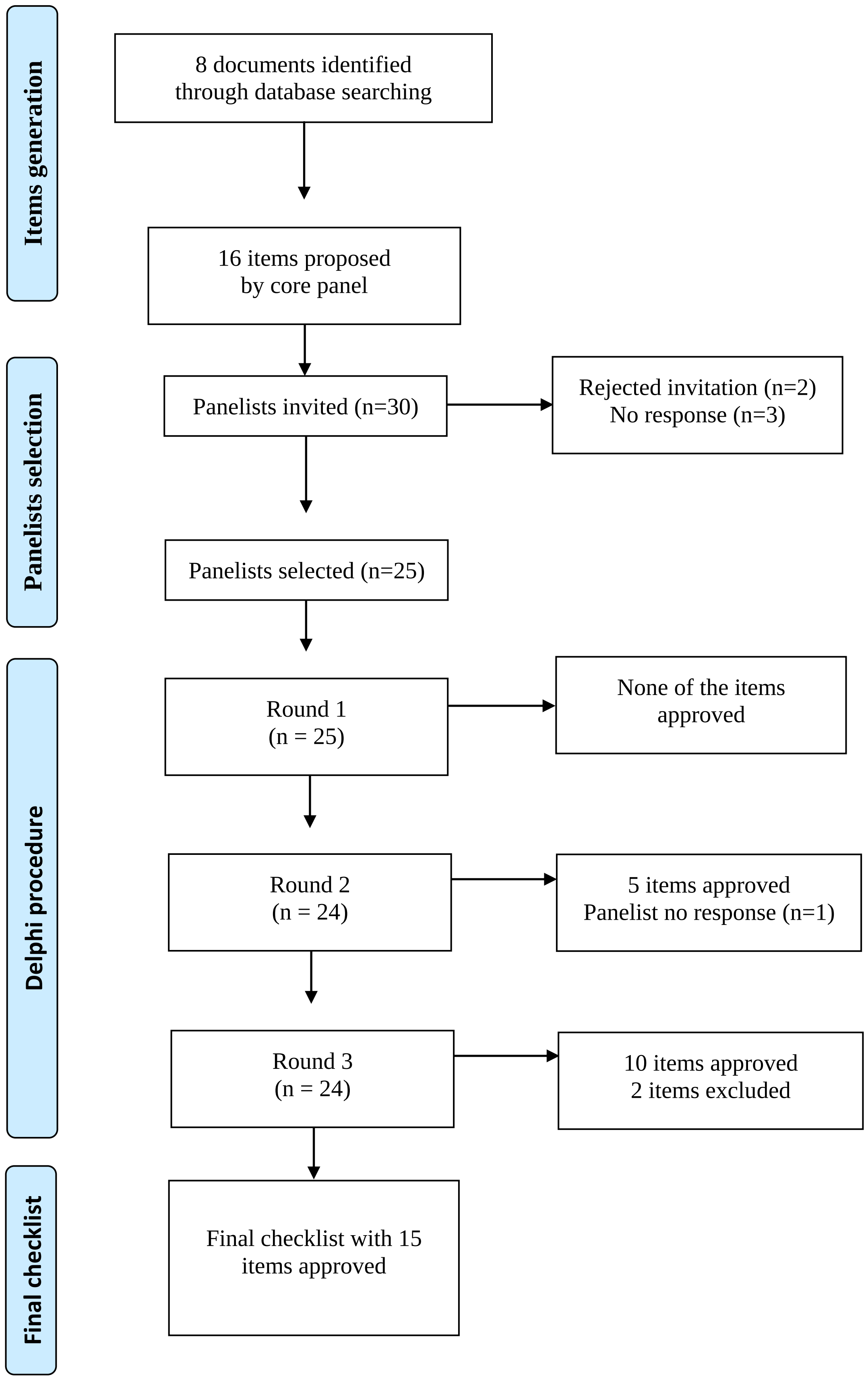

Figure 1. Flowchart of the development process (n = number of panelists).

Table 1. Panelists' responses through the rounds with core panel decisions for each item.

Item	Round 1 (n=25)						Round 2 (n=24)						Round 3 (n=24)					
	SA	MA	N	MD	SD	SA+MA: Decision	SA	MA	N	MD	SD	SA+MA: Decision	SA	MA	N	MD	SD	SA+MA: Decision
Individual data[1]	17 (68)	4 (16)	2 (8)	2 (8)	0 (0)	21 (84): modify	19 (79)	4 (17)	1 (4)	0 (0)	0 (0)	23 (96): modify	20 (83)	4 (17)	0 (0)	0 (0)	0 (0)	24 (100): approve
Previous instructions[2]	19 (76)	4 (16)	1 (4)	1 (4)	0 (0)	23 (92): modify	22 (92)	2 (8)	0 (0)	0 (0)	0 (0)	24 (100): rephrase	19 (79)	4 (17	1 (4)	0 (0)	0 (0)	23 (96): approve
Extrinsic factors[3]	22 (88)	2 (8)	1 (4)	0 (0)	0 (0)	24 (96): rephrase	22 (92)	1 (4)	1 (4)	0 (0)	0 (0)	23 (96): approve	a	a	a	a	a	a
Environmental condition[4]	12 (48)	4 (16)	2 (8)	4 (16)	3 (12)	16 (64): modify	20 (83)	3 (13)	1 (4)	0 (0)	0 (0)	23 (96) approve	a	a	a	a	a	a
Environmental setup[5]	18 (72)	7 (28)	0 (0)	0 (0)	0 (0)	25 (100): rephrase	19 (79)	4 (17)	0 (0)	1 (4)	0 (0)	23 (96): rephrase	18 (75)	2 (8)	1 (4)	3 (13)	0 (0)	20 (83): approve
Equipment[6]	13 (52)	11 (44)	0 (0)	1 (4)	0 (0)	24 (96): rephrase	19 (79)	5 (21)	0 (0)	0 (0)	0 (0)	24 (100): rephrase	15 (63)	7 (29)	0 (0)	2 (8)	0 (0)	22 (92): approve
Image background	11 (44)	7 (28)	3(12)	3 (12)	1(4)	18 (72): modify	17 (71)	2 (8)	4 (17)	1 (4)	0 (0)	19 (79): modify	13 (54)	3 (13)	6 (25)	2 (8)	0 (0)	16 (67): exclude
Aclimation[7]	12 (48)	8 (32)	0 (0)	3 (12)	2 (8)	20 (80): modify	16 (67)	5 (21)	1 (4)	2 (8)	0 (0)	21 (88): rephrase	16 (67)	4 (17)	1 (4)	3 (13)	0 (0)	20 (83): approve
Camera preparation[8]	15 (60)	6 (24)	1 (4)	2 (8)	1 (4)	21 (84): modify	19 (79)	3 (13)	1 (4)	1 (4)	0 (0)	22 (92): rephrase	17 (71)	4 (17)	0 (0)	3 (13)	0 (0)	21 (88): approve
Image recording[9]	13 (52)	8 (32)	2 (8)	2 (8)	0 (0)	21 (84): rephrase	19 (79)	3 (13)	0 (0)	2 (8)	0 (0)	22 (92): modify	16 (67)	6 (25)	0 (0)	2 (8)	0 (0)	22 (92) approve
Camera position[10]	22 (88)	2 (8)	1 (4)	0 (0)	0 (0)	24 (96): rephrase	23 (96)	1 (4)	0 (0)	0 (0)	0 (0)	24 (100): approve	a	a	a	a	a	a
Emissivity[11]	17 (68)	5 (20)	2 (8)	1 (4)	0 (0)	22 (88): rephrase	23 (96)	1 (4)	0 (0)	0 (0)	0 (0)	24 (100): approve	a	a	a	a	a	a
Assessment time[12]	12 (48)	8 (32)	1 (4)	3 (12)	1 (4)	20 (80): exclude#	9 (38)	5 (21)	3 (13)	4 (17)	3 (13)	14 (58): divide	17 (71)	4 (17)	2 (8)	1 (4)	0 (0)	21 (88): approve
Assessment operators*	-	-	-	-	-	-	-	-	-	-	-	-	12 (50)	5 (21)	6 (25)	0 (0)	1 (4)	17 (71): exclude
Body position[13]	15 (60)	6 (24)	2 (8)	1 (4)	1 (4)	21 (84): rephrase	18 (75)	5 (21)	0 (0)	1 (4)	0 (0)	23 (96): rephrase	17 (71)	6 (25)	0 (0)	1 (4)	0 (0)	23 (96): approve
Method of drying the skin[14]	16 (64)	5 (20)	2 (8)	1 (4)	1 (4)	21 (84): exclude#	13 (54)	4 (17)	5 (21)	1 (4)	1 (4)	17 (71): modify	19 (79)	3 (13)	2 (8)	0 (0)	0 (0)	22 (92): approve
Image evaluation[15]	22 (88)	1 (4)	2 (8)	0 (0)	0 (0)	23 (92): rephrase	24 (100)	0 (0)	0 (0)	0 (0)	0 (0)	24 (100): approve	a	a	a	a	a	a

Data expressed as number (percentage of responses); Superscript numbers refer to final checklist (see table 2); a = item approved in the previous round; * The item was proposed by the division of item "assessment time" in the third round; # The item was presented again in a different approach; SA = Strongly agree; MA = Moderately agree; N = Neutral; MD = Moderately disagree; SD = Strongly disagree; SA+MD = sum of the responses of strongly agree and moderately agree.

Table 2. Thermographic imaging in sports and exercise medicine (TISEM).

1) The relevant individual data of the participants must be provided. Note: These could include, but are not limited to, age, sex, body mass, height, body mass index, ethnicity and whether they are smokers or not. An indication of physical activity profile (e.g. frequency, duration, intensity, and activity description) should be reported.
❑ Yes ❑ No ❑ Unclear
2) Participants should be instructed to avoid alcohol beverages, smoking, caffeine, large meals, ointments, cosmetics and showering for four hours before the assessment. Also, sunbathing (e.g. UV sessions or direct sun without protection) should be avoided before the assessment. Note: This should be confirmed verbally before the assessment. The use of any medicinal treatments or drugs should be recorded. Any condition that could not be avoided should be reported.
❑ Yes ❑ No ❑ Unclear
3) Extrinsic factors affecting skin temperature (e.g. physical activity prior to the assessment, massage, electrotherapy, ultrasound, heat or cold exposure, cryotherapy) should be clearly described.
❑ Yes ❑ No ❑ Unclear
4) Ambient temperature and relative humidity of the location where the assessment took place must be recorded and reported as mean ± standard deviation.
❑ Yes ❑ No ❑ Unclear
5) The assessment should be completed away from any source of infrared radiation (e.g. electronic devices, lightning) or airflow (e.g. under an air conditioning unit). Note: Any condition that could not be controlled should be reported.
❑ Yes ❑ No ❑ Unclear
6) The manufacturer, model and accuracy of the camera used should be provided. Note: When available it is recommended to provide the maintenance information of the equipment (e.g. when and where it was completed the last calibration).
❑ Yes ❑ No ❑ Unclear
7) An acclimation period in the examination room should be completed. Note: This item is only applicable for initial baseline measurements or basal analysis.
❑ Yes ❑ No ❑ Unclear
8) If necessary the camera should be turned on for some time prior to the test to allow sensor stabilization following the manufacturer's guidelines.
❑ Yes ❑ No ❑ Unclear
9) Conditions of image recording such as mean distance between object and camera, percentage of the region of interest within the image should be detailed.
❑ Yes ❑ No ❑ Unclear
10) The camera should be positioned perpendicular to the region of interest.
❑ Yes ❑ No ❑ Unclear
11) Emissivity settings of the camera must be reported. Note: 0.98 of emissivity is suggested for a dry clean skin surface.
❑ Yes ❑ No ❑ Unclear
12) The time of day at which the images were taken should be reported.
❑ Yes ❑ No ❑ Unclear
13) The standard body position of the subject and the regions of interest must be well described and appropriately selected. A visual example (with temperature scale presented and scale of colors properly configured) is recommended.
❑ Yes ❑ No ❑ Unclear
14) If the skin is dried (e.g. to remove surface water), the drying method should be clearly described.
❑ Yes ❑ No ❑ Unclear
15) The evaluation of thermograms and collection of temperature from the software should be clearly described.
❑ Yes ❑ No ❑ Unclear