Clinical accuracy of infrared temperature measurement devices: a comparison against non-invasive core-body temperature

Authors: Thomas Holder,[A] Frances Sophie Woodley Hooper,[B] David Yates,[C] Zion Tse,[D] Samadhan Patil,[E] Ahmed Moussa,[A] Lucy Batten,[A] Vignesh Radhakrishnan,[B] Mark Allison,[F] Catherine Hewitt,[G] Ada Keding,[H] Greg Forshaw[I] and Vijay Jayagopal[J]

ABSTRACT

During the coronavirus 2019 (COVID-19) pandemic, the implementation of non-contact infrared thermometry (NCIT) became an increasingly popular method of screening body temperature. However, data on the accuracy of these devices and the standardisation of their use are limited. In the current study, the body temperature of non-febrile volunteers was measured using infrared (IR) thermography, IR tympanic thermometry and IR gun thermometry at different facial feature locations and distances and compared with SpotOn core-body temperature. Poor agreement was found between all IR devices and SpotOn measurements (intra-class correlation coefficient <0.8). Bland–Alman analysis showed the narrowest limits of agreement with the IR gun at 3 cm from the forehead (bias = 0.19˚C, limits of agreement (LOA): –0.58˚C to 0.97˚C) and widest with the IR gun at the nose (bias = 1.40˚C, LOA: –1.15˚C to 3.94˚C). Thus, our findings challenge the established use of IR thermometry devices within hospital settings without adequate standard operating procedures to reduce operator error.

KEYWORDS: infrared thermometry, COVID-19, medical devices, temperature

DOI: 10.7861/clinmed.2022-0252

Introduction

Accurate temperature measurement is a key assessment during the early identification of sepsis and evaluation of infections, such as coronvirus 2019 (COVID-19). A temperature error of 0.2˚C can have major clinical implications in a healthcare setting,[1] including decisions on antimicrobial therapy, drawing blood cultures and patient isolation. During the COVID-19 pandemic, temperature screening became commonplace because one of the most prevalent diagnostic indications of COVID-19 is an elevated temperature of ≥37.8˚C. Around 78% of cases of COVID-19 present with a fever.[2]

Clinical methods of measuring temperature include invasive, contact and non-contact. Invasive devices, such as rectal,[3] pulmonary artery, oesophageal and urinary bladder thermometers,[4] are the most accurate and are often used as reference temperatures when evaluating the accuracy of contact and non-contact devices.[3,5–10] However, invasive methods are unsuitable for triage and mass temperature screening. By contrast, contact methods, such as axillary and tympanic thermometers, offer faster measurement times and reduced discomfort to patients. Nevertheless, they are not as accurate as invasive methods[4,6,11–13] and still require high levels of staff–patient contact. Additionally, tympanic thermometers have a high level of waste associated with them because of their disposable hygiene covers.[14]

Non-contact methods of temperature measurements, such as infrared (IR) thermometers and thermal cameras, reduce contact time with potentially infected patients and offer rapid measurements for mass temperature screening. The COVID-19 pandemic has catalysed a widespread use of non-contact IR thermometry (NCIT) for fever detection.

Despite their increased popularity and use during the pandemic, little research has been conducted into NCIT device accuracy or application methods. It is known that IR thermometry devices have decreasing accuracy with increasing distance.[15] It is standard practice to measure facial temperature with a NCIT device, but there is no uniformity or evidence on which location is optimal for this purpose. Device manufacturers state that NCIT forehead devices should be used within a limited working distance from the forehead only, with no obstruction by hair or sweat.[16] However, some studies have evaluated these guidelines and have found that NCIT wrist measurements are more stable and accurate compared with forehead measurements[11,13]

Sjoding *et al*[17] recently shed light on racial bias with the use of pulse oximetry, which highlights potential additional factors that might influence the accurate measurement of vital signs using indirect methods, such as NCIT. It is known that differing amounts of solar energy are absorbed by different colours, which could also

Authors: [A]trust grade registrar, York and Scarborough Teaching Hospital NHS Foundation Trust, York, UK; [B]PhD research assistant, University of York, York, UK; [C]consultant in anaesthetics, York and Scarborough Teaching Hospital NHS Foundation Trust, York, UK; [D]professor and chair in medical robotics, Medical Robotics Lab, University of York, York, UK; [E]lecturer in medical engineering, University of York, York, UK; [F]medical student, University of York, York, UK; [G]professor of trials and statistics, University of York, York, UK; [H]lecturer and researcher of trials and statistics, University of York, York, UK; [I]clinical trials manager, York Teaching Hospital NHS Foundation Trust, York, UK; [J]consultant in diabetes and endocrinology, York and Scarborough Teaching Hospital NHS Foundation Trust, York, UK

lead to discrepancies in IR emission and potential inaccuracies when using NCIT devices.[18]

Therefore, the current study was performed to assess the accuracy of three non-invasive devices: a non-contact IR forehead thermometer (IR gun), a tympanic thermometer and an IR thermal camera. These three devices were compared to a gold-standard for measuring core-body temperature non-invasively, the 3M Bair Hugger Temperature Monitoring System (SpotOn).[9,19,20] To investigate the manufacturer's guidelines, the IR gun was used at varying facial locations and distances. To evaluate the current use of NCIT devices in clinical settings, a survey was created and distributed to nurses at a hospital trust. The aim of this study was to establish best-practice parameters for the site and distance from skin, to optimise the use of NCIT devices.

Methods

Equipment

Four different temperature measurement methods and devices were used in this study: (i) SpotOn (3M, USA); (ii) the JXB-182 Infrared Forehead Thermometer (IR gun) (Berrcom, China); (iii) the Genius 3 Tympanic Thermometer (tympanic thermometer) (Covidien, USA); and (iv) the FLIR C3 Thermal Camera (IR camera) (FLIR, USA).

A rig was used to mount an IR gun to standardise its distance from each participant's face during stabilised measurements. This allowed investigation of distance and location effects on IR gun accuracy by incorporating three standard IR gun distance settings and a chin rest to stabilise the head. The rig also enabled us to accurately replicate the manufacturers' guidelines, which state a measurement distance of 3–5 cm taken from the forehead.[16]

Environmental controls and participant selection

Environmental controls, including thermostat temperature (20°C), closed windows, and drawn blinds, were set before data acquisition and checked routinely to reduce their effect on participant temperature readings.

In total, 120 non-febrile adult volunteers (aged 20 to 59 years) were recruited to the study, with written informed consent and confirmation of eligibility forms completed by each participant. Exclusion criteria included skin disorders and presenting with symptoms indicative of COVID-19. Participants were healthcare professionals, administration staff, medical students and the research team working with York Teaching Hospital NHS Trust. Eligible participants were asked to remove glasses or face coverings and the room temperature was measured using a thermometer probe (±0.1°C precision, ±0.5°C accuracy) before participant temperature measurements. There were five investigators, each having the same role throughout. Two investigators gained consent and completed eligibility forms, one performed thermal imaging and two carried out the remaining temperature measurements. The data collection took place over 4 days.

IR camera workstation

First, IR and digital photos were taken using a FLIR C3 thermal camera and a Moto G8 smartphone. The tripod-mounted IR camera was positioned 0.6 m from the participant's face and the participants were positioned to align their head with a reference frame to maintain image ratios (Fig 1). One IR image was taken of each participant in the coronal plane for post-processing analysis and one digital photo was taken for participant reference.

Fig 1. The FLIR C3 IR camera workstation with participant alignment, digital photography and IR thermography. IR = infrared.

SpotOn

Following IR photography, a Bair Hugger disposable temperature sensor was fitted to the participant's left temple, avoiding their hairline to ensure optimal adhesion. The sensor was connected to the Bair Hugger control unit to allow for equilibrium to be reached while IR and tympanic temperature measurements were recorded. Once calibrated, core-body temperature from the SpotOn system was recorded.

IR forehead thermometer

Nine IR gun temperature measurements were taken of each participant, including free-hand measurements of their facial features (eyes, nose and lips) at 3–5 cm, and stabilised measurements of their forehead and right temple. The stabilised measurements were taken by placing the participant's chin upon the chin rest of the wooden rig, facing the mounted IR gun for forehead measurements and turning their head for temple measurements. Their temperature was then recorded at set measured distances (3, 6 and 9 cm) by translating the mounted IR thermometer backward. Stabilised measurements were then repeated if the participant consented to remove any cosmetics present from their forehead and temple.

Tympanic thermometer

Tympanic temperatures of both ears of each participant were recorded once each. Thermometer probe covers were disposed of between measurements.

Thermal imaging post-processing and analysis

Following data collection, thermal photos were processed and analysed using FLIR Tools (FLIR, USA). The parameters listed in Table 1 were implemented during post-processing

Table 1. FLIR tool parameters changed during thermal image post-processing

Parameter name	Parameter description	Change in parameter values
Distance	Distance between camera and participant	0.6 m
Atmospheric temperature	Room temperature recorded by thermometer probe before data acquisition from each participant	Record range: 19.6–21.9°C
Reflective temperature	Thermal radiation of background objects; assumed to be the same as atmospheric temperature	Record range: 19.6–21.9°C
Thermal tuning	Adjustment of level and span of thermal aspect of image to highlight regions with temperature within desired range	Maximum temperature: 39°C Minimum temperature: 30°C

to match the environmental conditions, and the maximum temperature and its location for each participant was established (Fig 2).

Survey methodology

A survey of 93 healthcare professionals who regularly used IR devices was undertaken to determine typical measurement techniques. The candidates were asked to answer questions based on the distance from, and where on, the person the measurement was made. The survey results were compared with the manufacturers' guidelines (3–5 cm from the forehead) and the percentage of compliance with the guidelines was assessed. The survey also assessed the amount, and formality, of training that staff received on the use of IR devices. Supplementary material S1 details the questionnaire used.

Sample size

Assuming a standard deviation (SD) of measurement differences of ~0.5°C and a standard error (SE) around the limit of agreement of 0.32,[21] we calculated that a sample size of 110 would be needed to provide 80% power to assess agreement between two methods of temperature measurement with a mean difference of zero.[22] Even though a measurement difference of 0.2°C would be clinically significant, an agreement difference of 0.5°C was chosen because it was felt that this would demonstrate a difference between devices more meaningfully.

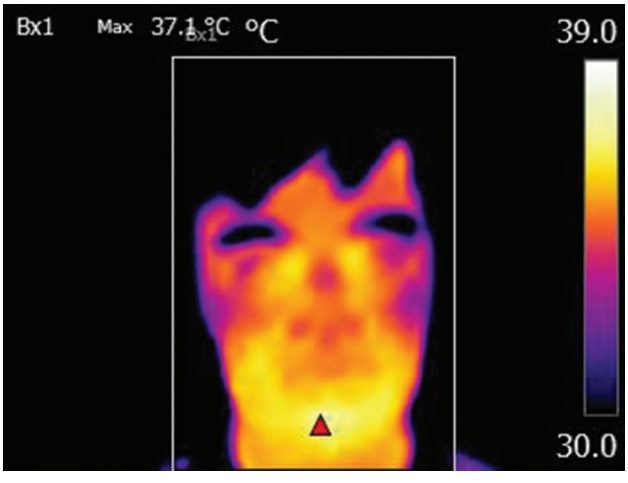

Fig 2. IR thermal image of head and neck taken by the FLIR C3 camera after post-processing with maximum temperature and location. IR = infrared.

Data analysis

Stata version 16 was used for all statistical analyses. Agreement between temperature measurements for the same individuals was assessed using intra-class correlation coefficients (ICCs) between SpotOn measurements as the reference value and each of the other devices. ICCs were calculated from two-way mixed effects models using the kappaetc command. We used the following categories to assess reliability: values less than 0.5 were indicative of poor reliability; values between 0.5 and 0.75 indicated moderate reliability; values between 0.75 and 0.9 indicated good reliability and values greater than 0.90 indicated excellent reliability.[23] Device readings were further compared with the SpotOn temperature using Bland–Altman analysis, assessing the magnitude and direction of mean differences as well as the width of the limits of agreement (LOA). Agreement within the same instrument was assessed using ICCs for IR gun by location (eyes, nose and lips) and target area (forehead or temple at different distances) and for tympanic measurement (left and right). One volunteer was excluded from the study because of early withdrawal resulting in missing temperature data.

To evaluate the effect of ethnicity on NCIT accuracy, further Bland–Altman analyses were performed on each of the racial groups populating the study. Mean temperature differences between the SpotOn and IR camera and IR Gun at the forehead and temple at 3 cm were compared, with LOA, between the six ethnic groups: White, Black, East Asian, West Asian, South Asian and mixed.

The study was funded by an unrestricted educational grant from the Elsie May Sykes Trust and was sponsored by York Teaching Hospitals NHS Trust. Ethical approval was provided by the Health Research Authority.

Results

Demographic data

In total, 119 volunteers participated (50 males, 69 females) with a median age of 28 (SD: 11) years. No participants presented with an elevated temperature (SpotOn ≥37.8°C). All 119 participants completed the assessments.

Survey results

Of the 93 staff surveyed, 32 (34%) reported that they had received training, ranging from informal observational training to formal induction sessions, on how to use IR guns. No staff measured temperatures outside of the forehead or temple, with 22.6% of staff measuring the forehead, 75.3% measuring the temple, and 2% measuring both. In terms of distance, 50.5% followed the manufacturer's guidance (3–5 cm), and 36.6% measured within 1 cm of the skin surface.

Maximum temperature locations (IR images)

Figure 3 shows the maximum temperature location of each participant against the frequency at which they occurred. Maximum temperature was located within the eye region of 68 (57.1%) participants, the mouth region of 23 (19.3%) participants and the neck region of 22 (18.5%) participants.

Intra-class corelation coefficients

The ICC with 95% confidence interval (CI) results are shown in Table 2, representing the agreement in temperature measurements between different locations and distances with the same device. Table 2 additionally shows the agreement in temperature measurements between the SpotOn and all other devices at all locations and distances, along with their mean differences and LOAs. No acceptable agreement (ICC >0.8) was shown between any sets of measurements. Temple and nose measurements showed approximately zero agreement with the SpotOn. Best agreement with SpotOn was achieved with tympanic left and right readings (ICC = 0.36 and 0.37, respectively). Within the same target area (forehead or temple), measurements showed moderate correlation between different distances (3 cm, 6 cm and 9 cm).

Bland–Altman analysis

Based on the mean difference between readings (Table 2), the average bias was for SpotOn to measure higher temperatures than any other device, with LOAs ranging between –1.15 and 3.94°C. Compared with SpotOn, the smallest mean difference was from the IR gun at the eye (0.07°C, LOA –0.88-1.02°C) and the narrowest LOAs were for the forehead measurement at 3 cm (–0.58-0.97°C, mean 0.19°C). For the IR Gun at 3 cm from the forehead, participants with the lowest mean temperature error were of West Asian descent (0.20°C, LOA: –0.41–1.0°C), with those identifying as mixed race having the largest mean temperature difference (0.55°C) with the smallest LOA range (0.41–0.69°C). Similar results were seen with the IR gun measurements at the forehead.

Discussion

This study was unable to demonstrate acceptable agreement between any sets of temperature measurement and found a high degree of variation in recorded temperature dependent on the instrument used. These findings challenge the hitherto commonly unquestioned acceptance in traditional medical practice that a temperature measurement reflects an accurate measure.

The manufacturer's guidelines for the IR gun state that measurements should be taken 3–5 cm from the patient's forehead.[16] Our methods explored the accuracy of the device in terms of the results, both following these guidelines and not. When compared with SpotOn, IR temperatures measured at 3 cm from the participant's forehead gave the narrowest LOA (–0.58-0.97°C) compared with all other devices, with a low mean difference of 0.19°C. This narrow LOA exceeded the manufacturer's stated accuracy of ±0.2°C,[16] and poor agreement was found between SpotOn and the IR gun at this location (ICC = 0.32). Bijur *et al*[6] found similar results; a low mean difference of 0.0°C with wide LOA (–1.1-1.1°C) when comparing IR temporal artery and rectal temperatures. Allegaert *et al*[7] also compared rectal temperatures to IR skin measurements in children, finding a significant difference (p=0.001) between device measurements, with a mean difference of 0.34°C and wide LOA (–0.92-1.60°C).

Several studies compared NCIT devices with tympanic thermometry. Ng *et al*[12] studied the agreement and correlation between tympanic and NCIT forehead measurements in children, finding a large mean error between the devices (2.34°C, LOA: 0.26–4.42°C) and a negative correlation between mean error

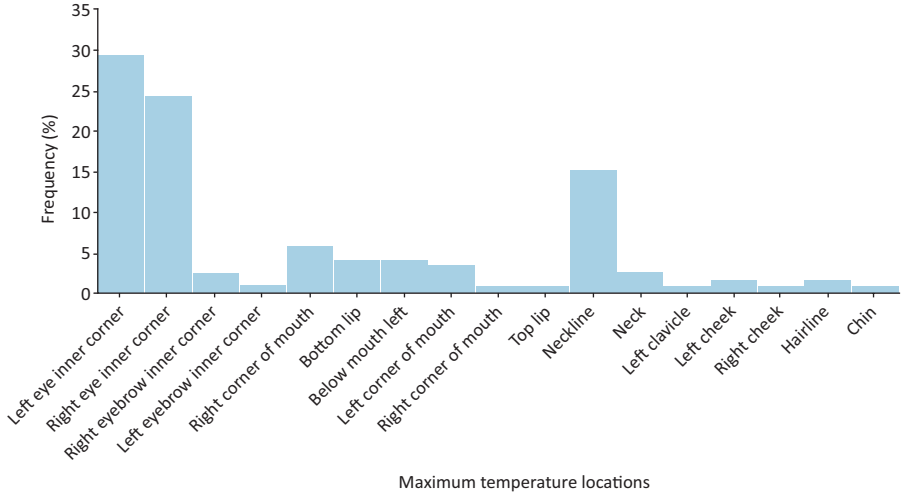

Fig 3. Occurrence frequency of maximum facial temperature locations from the IR images. IR = infrared.

Table 2. Agreement between the different locations and distances of the IR temperature measurement devices and between the SpotOn and other IR temperature measurement devices

IR device	Mean (SD)	Mean difference (LOA)[a]	ICC (95% CI)
IR gun			
Location: eyes, nose and lips	–	–	0.15 (0.04, 0.27)
Distance head: 3, 6 and 9 cm	–	–	0.71 (0.63, 0.78)
Distance temple: 3, 6 and 9 cm	–	–	0.72 (0.65, 0.79)
Tympanic thermometer			
Left and right ear	–	–	0.66 (0.54, 0.75)
SpotOn versus IR device			
SpotOn	36.9 (0.37)	–	–
IR gun location			
Eyes	36.9 (0.36)	0.07 (–0.88, 1.02)	0.13 (0.00, 0.30)
Nose	35.5 (1.26)	1.40 (–1.15, 3.94)	0.00 (0.00, 0.20)
Lips	36.5 (0.38)	0.41 (–0.49, 1.31)	0.25 (0.08, 0.41)
IR gun distance			
Forehead: 3 cm	36.7 (0.30)	0.19 (–0.58, 0.97)	0.32 (0.14, 0.47)
Forehead: 6 cm	36.6 (0.22)	0.35 (–0.43, 1.14)	0.15 (0.00, 0.32)
Forehead: 9 cm	36.4 (0.23)	0.51 (–0.32, 1.35)	0.07 (0.00, 0.24)
Temple: 3 cm	36.7 (0.48)	0.24 (–0.93, 1.41)	0.05 (0.00, 0.22)
Temple: 6 cm	36.5 (0.42)	0.42 (–0.65, 1.49)	0.05 (0.00, 0.23)
Temple: 9 cm	36.4 (0.46)	0.57 (–0.60, 1.73)	0.00 (0.00, 0.18)
Tympanic			
Left	36.3 (0.49)	0.59 (–0.37, 1.55)	0.36 (0.19, 0.51)
Right	36.3 (0.49)	0.64 (–0.31, 1.60)	0.37 (0.20, 0.51)
IR image			
Coronal plane	36.4 (0.53)	0.50 (–0.58, 1.59)	0.27 (0.10, 0.43)

[a]SpotOn minus comparator measurement.

CI = confidence interval; ICC = intra-class correlation coefficient; IR = infrared.

and average temperature measurement, rejecting the clinical use of NCIT for fever screening. Chen *et al*[13] also compared NCIT and tympanic measurements. In an indoor setting, wide LOAs were found (–1.91–0.80°C), with 57.1% of measurements having a clinically acceptable error of ±0.5°C. Although these studies indicate inaccuracies between NCIT devices and tympanic measurements, our results suggest that tympanic thermometers should not be used as a reference for core-body temperatures. Despite tympanic measurements having the highest agreement with SpotOn, agreement was still below acceptable levels (ICC <0.8). Large mean differences between the two devices were also observed (left: 0.59°C, right: 0.64°C), with the LOA (left: –0.37–1.55°C, right: –0.31–1.60°C) exceeded the manufacturer's stated accuracy of ±0.3°C.[14] Bijur *et al*[6] found similar inaccuracies when comparing tympanic temperatures with rectal core-body temperature, with a mean difference of –0.3°C and LOA range of –1.2–0.6°C. These inaccuracies between tympanic and core-body temperatures result from cerumen, hair and a narrow canal circumference obstructing the tympanic membrane and thermometer probe.[5,24]

Given that the IR gun had a maximum working distance of 5 cm, it was expected that increasing its distance from the forehead and temple would reduce the result accuracy. At corresponding distances between the temple and forehead, temple measurements showed lower agreement and higher mean difference and LOAs between the IR gun and SpotOn. Similarly, Fong *et al.* found a 0.1°C mean error between temporal and forehead IR measurements, with the temple measuring lower than the forehead on average and with LOA (–0.8–0.7°C) outside the acceptable range for the study (±0.5°C).[25] These variations might result from hair obstructing the thermometer at the temple[16,25,26] and, thus, support the manufacturer's instruction for forehead use only.

The NCIT results showed reduced variation when using the devices as recommended by the manufacturer. However, our results suggest no standardisation of NCIT device usage based on these recommendations by healthcare staff, with only 22% of staff using the devices at the recommended forehead location. Given a variation in NCIT device manufacturers used by different departments and hospitals, it is imperative that the

instructions for each device on optimal measurement location and distance are followed so as to not alter measurement accuracy. Alternatively, homogeneity of devices within the NHS would minimise variation and allow for a NCIT device standardised operating procedure, including staff training and standardised measurement parameters. These would include a set measurement distance of 5 cm, visualised as the width of a credit card, from the forehead, and the removal of obstructions, such as hair. Additionally, skin should be cleansed of cosmetics and sweat and dried well before temperatures are taken. Given the inaccuracies of NCIT devices found during this study, the device limitations must be considered clinically when interpreting temperature results.

Most maximum temperatures were found in the regions of the inner eye (57.1%), mouth (19.3%) and neck (18.5%), rather than the traditionally measured forehead location (1.7%). Zhou *et al*[27] found comparable distributions of maximum facial temperatures in the region of the inner eye (59.5%). This suggests the inner eye as an alternative location to traditional IR forehead measurements, which is supported by our results comparing SpotOn and the IR gun at the eyes, which gave the lowest mean temperature difference of any other comparison (0.07°C). However, Fernandes *et al*[28] compared IR thermography at the inner canthus to core-body temperature from an ingestible thermal pill and found poor agreement (ICC = 0.49), with wide LOA (–1.275-0.049°C).

Limitations and future work

Temperature measurement in a patient and the subsequent clinical decision making that follows can have profound effects for an individual and for health and social care systems. Therefore, a paradigm shift in our understanding of the devices used and a more nuanced approach to how we use the output in clinical practice is required.

Several demographic factors restricted the depth of our investigation. With an ethnically imbalanced cohort, in which 83% of participants were of White British background, we were unable to fully assess the effects of skin tone and ethnicity on IR thermometer measurements. Although there was a trend toward skin colour being a variable in IR temperatures readings, our cohort was not diverse enough to provide statistically significant results. Additionally, investigations into the effect of cosmetics were limited because of low volumes of participants volunteering to remove cosmetics for additional temperature measurements. Our results were also limited by studying only non-febrile participants. Future investigations will focus on a more ethnically diverse cohort and in-depth studies on cosmetics effects on NCIT device accuracy, as well as the ability of IR thermometers to detect known fevers. The study was not powered to provide precision for the ICCs because the target was the agreement based on the mean difference, which limits the conclusions that can be drawn. SpotOn measurements were used as the reference value in the analysis and, because all other methods systematically underestimated the 'true' temperature, this might indicate that SpotOn is a consistent gold standard. Conversely, it could question the accuracy of SpotOn and, thus, a confirmatory measure might be warranted.

In conclusion, NCIT device availability offered quicker and more convenient ways to measure patient and customer temperatures throughout the UK during the COVID-19 pandemic. However, the ability of NCIT devices to reflect accurate core-body temperature appears limited. When NCIT devices are used, particular care is needed to ensure that staff using the devices are appropriately trained and that all staff are made aware of device limitations. We need to consider refining our use of temperature measurements from an absolute value to thinking more in terms of change from baseline, to trigger subsequent clinical action. However, more research is urgently required to help establish best practice here. ■

Summary

What is known?

Elevated temperature is a key indicator of infectious disease and aids with planning clinical decisions, such as performing testing, treatment plans and maintaining patient isolation. The implementation of NCIT during the COVID-19 pandemic has offered a way for businesses to screen customers and, in hospital, for nurses to reduce contact with potentially infectious patients. Despite their increased popularity and widespread use, little research has been conducted into the accuracy of NCIT devices and the standardisation of methods in which they are used.

What is the question?

What is the current usage of IR thermometry devices (IR gun thermometer, IR tympanic thermometer and IR thermal camera) within hospital settings? What is the clinical accuracy of these devices for screening core-body temperature? What standard operating procedure should be suggested for using IR gun thermometers within a hospital setting?

What was found?

Compared with SpotOn, all IR devices showed poor agreement in core-body temperature measurements, with poor accuracy and repeatability of their measurements.

What is the implication for practice now?

Hospitals and other sectors should not rely on IR gun thermometers to accurately screen for core-body temperature; if their use were to continue, ensuring that the device is used 3–5 cm away from a clean, dry forehead should be standard practice.

Supplementary material

Additional supplementary material may be found in the online version of this article at www.rcpjournals.org/clinmedicine:
S1 – Questionnaire.

Conflicts of interests

Grant funding from the York Teaching Hospitals NHS Foundation Trust Elsie May Sykes Award provided a stipend for five of the authors; FSWH and VR as PhD research assistants, and TH, LB and AM as Trust-grade registrar research assistants.

Acknowledgements

We thank Deborah Phillips (York Teaching Hospital NHS Foundation Trust) for her guidance and contributions toward the development of our research; we thank Michael Li (University of York) for his help with temperature data collection and post-possessing image analysis.

References

1 Farnell S, Maxwell L, Tan S, Rhodes A, Philips B. Temperature measurement: comparison of non-invasive methods used in adult critical care. *J Clin Nurs* 2005;14:632–9.
2 Grant MC, Geoghegan L, Arbyn M *et al*. The prevalence of symptoms in 24,410 adults infected by the novel coronavirus (SARS-CoV-2; COVID-19): A systematic review and meta-analysis of 148 studies from 9 countries. *PLoS ONE* 2020;15:e0234765.
3 Walker GA, Runde D, Rolston DM, Wiener D, Lee J. Emergency department rectal temperatures in over 10 years: a retrospective observational study. *World J Emerg Med* 2013;4:107–12.
4 McCallum L, Higgins D. Measuring body temperature. *Nurs Times* 2012;108:20–2.
5 Daanen HA. Infrared tympanic temperature and ear canal morphology. *J Med Eng Technol* 2006;30:224–34.
6 Bijur PE, Shah PD, Esses D. Temperature measurement in the adult emergency department: oral, tympanic membrane and temporal artery temperatures versus rectal temperature. *Emerg Med J* 2016;33:843–7.
7 Allegaert K, Casteels K, van Gorp I, Bogaert G. Tympanic, infrared skin, and temporal artery scan thermometers compared with rectal measurement in children: a real-life assessment. *Curr Ther Res Clin Exp* 2014;76:34–8.
8 Eshraghi Y, Nasr V, Parra-Sanchez I *et al*. An evaluation of a zero-heat-flux cutaneous thermometer in cardiac surgical patients. *Anesth Analg* 2014;119:543–9.
9 Makinen MT, Pesonen A, Jousela I *et al*. Novel zero-heat-flux deep body temperature measurement in lower extremity vascular and cardiac surgery. *J Cardiothorac Vasc Anesth* 2016;30:973–8.
10 Kollmann Camaiora A, Brogly N, Alsina E, de Celis I, Huercio I, Gilsanz F. Validation of the Zero-Heat-Flux thermometer (SpotOn(R)) in major gynecological surgery to monitor intraoperative core temperature: a comparative study with esophageal core temperature. *Minerva Anestesiol* 2019;85:351–7.
11 Chen HY, Chen A, Chen C. Investigation of the impact of infrared sensors on core body temperature monitoring by comparing measurement sites. *Sensors (Basel)* 2020;20:2885.
12 Ng DK, Chan CH, Lee RS, Leung LC. Non-contact infrared thermometry temperature measurement for screening fever in children. *Ann Trop Paediatr* 2005;25:267–75.
13 Chen G, Xie J, Dai G *et al*. Validity of the use of wrist and forehead temperatures in screening the general population for COVID-19: a prospective real-world study. *Iran J Public Health* 2020;49(Suppl 1): 57–66.
14 Covidien. GeniusTM 3 Tymapnic Thermometer and Base Operating Manual 2016. www.cardinalhealth.co.uk/content/dam/corp/products/professional-products/ous-patient-recovery/documents/cardinal-health-genius-3-user-manual-2.pdf [Accessed 24 February 2023]
15 Chen J, Wang J-P, Shen T-Y, Xiong D, Liquan G. High precision infrared temperature measurement system based on distance compensation. *ITM Web Conferences* 2017;12:03021.
16 Non-contact Infrared Thermometer Model: JXB-182 Instructions For Use. www.kasa.cz/document/6/4/9/doc_4520946.pdf [Accessed 24 February 2023].
17 Sjoding MW, Dickson RP, Iwashyna TJ, Gay SE, Valley TS. Racial Bias in Pulse Oximetry Measurement. *N Engl J Med* 2020;383: 2477–8.
18 Stuart-Fox D, Newton E, Clusella-Trullas S. Thermal consequences of colour and near-infrared reflectance. *Philos Trans R Soc Lond B Biol Sci* 2017;372:20160345.
19 Kato H, Kawashima S, Mimuro S, Obata Y, Doi M, Nakajima Y. An evaluation of deep-forehead temperature (spoton®) in ICU patients after cardiac surgery. *Intensive Care Med Exp* 2015;3:A111.
20 Iden T, Horn EP, Bein B, Bohm R, Beese J, Hocker J. Intraoperative temperature monitoring with zero heat flux technology (3M SpotOn sensor) in comparison with sublingual and nasopharyngeal temperature: an observational study. *Eur J Anaesthesiol* 2015;32:387–91.
21 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–10.
22 Lu MJ, Zhong WH, Liu YX, Miao HZ, Li YC, Ji MH. Sample size for assessing agreement between two methods of measurement by Bland-Altman method. *Int J Biostat* 2016;12:20150039.
23 Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016;15:155–63.
24 Twerenbold R, Zehnder A, Breidthardt T *et al*. Limitations of infrared ear temperature measurement in clinical practice. *Swiss Med Wkly* 2010;140:w13131.
25 Fong WWS, Yeo SK, Fook-Chong SMC, Phang JK, Sim E. Comparison of temperature readings using infrared thermometers at three different sites: tympanic, forehead and temporal. *Proc Singapore Healthc* 2021;30:41–3.
26 Mangat J, Standley T, Prevost A, Vasconcelos J, White P. A comparison of technologies used for estimation of body temperature. *Physiol Meas* 2010;31:1105–18.
27 Zhou Y, Ghassemi P, Chen M *et al*. Clinical evaluation of fever-screening thermography: impact of consensus guidelines and facial measurement location. *J Biomed Opt* 2020;25:097002.
28 Fernandes AA, Moreira DG, Brito CJ *et al*. 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* 2016;62:50–5.

Address for correspondence: Zion Tse, Medical Robotics Lab, Department of Electronic Engineering, University of York, York, UK.
Email: zion.tse@york.ac.uk