



## Original Contributions

### Body Core Temperature Assessment in Emergency Care Departments

Hein A.M. Daanen, PHD,\* Gercora Hoitinga, MSc,<sup>†</sup> David J. Kruijt, BSc,<sup>†</sup> Kevin S. Koning, MSc,\*  
 Pooh P. Verheijen, BSc,\* Sanne I.M. de Baas, BSc,\* Anouk R. Bergsma, BSc,\* Cathelijne E. Snethlage, BSc,\*  
 Iman Al-Bander, BSc,\* and Lennart P.J. Teunissen, PHD<sup>‡</sup>

\*Department of Human Movement Sciences, Faculty of Behavioural and Movement Sciences, Amsterdam Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands, <sup>†</sup>Department of Emergency Medicine, Amsterdam UMC, University of Amsterdam, Meibergdreef, Amsterdam, The Netherlands, and <sup>‡</sup>Department of Training Medicine and Training Physiology, Royal Netherlands Army, Ministry of Defense, Utrecht, The Netherlands

Reprint Address: Hein A.M. Daanen, PhD, Department of Human Movement Sciences, Faculty of Behavioural and Movement Sciences, Amsterdam Movement Sciences, Vrije Universiteit Amsterdam, Van der Boechorststraat 7, 1081BT Amsterdam, The Netherlands

**□ Abstract—Background:** There is concern that the values provided by devices using infrared thermometry in emergency departments (EDs) do not reflect body core temperature accurately. **Objectives:** Evaluation of three thermometers commonly used in the ED. **Methods:** Two infrared ear thermometers and an infrared forehead thermometer were evaluated using 1) the Voltcraft IRS-350 calibration device, 2) comparing temperature values to a rectal end-exercise temperature (T-RECT) of 38.1°C in 12 participants, and 3) comparing temperature values to rectal temperature in 133 ED patients. **Results:** Calibration across the human core temperature range revealed that the ear thermometers underestimated radiant temperature by  $0.77 \pm 0.39^\circ\text{C}$  and  $1.84 \pm 0.26^\circ\text{C}$ , respectively, whereas the forehead thermometer overestimated radiant temperature by  $0.90 \pm 0.51^\circ\text{C}$ . After cycling exercise, all thermometers underestimated T-RECT ( $0.54 \pm 0.27^\circ\text{C}$  and  $1.03 \pm 0.48^\circ\text{C}$  for the ear thermometers and  $1.14 \pm 0.38^\circ\text{C}$  for the forehead thermometer). In the ED, the ear thermometers underestimated T-RECT by  $0.31 \pm 0.37^\circ\text{C}$  and  $0.46 \pm 0.50^\circ\text{C}$ , whereas the forehead thermometer exhibited a nonsignificant overestimation of  $0.04 \pm 0.46^\circ\text{C}$ . If the threshold for fever in all systems had been set to 37.5°C instead of 38.0°C, the sensitivity and specificity of the systems for real fever ( $T\text{-RECT} \geq 38^\circ\text{C}$ ) are, respectively, 71% and 96% (ear thermometer 1), 57% and 97% (ear thermometer 2), and 86% and 90% (forehead thermometer). **Conclusion:** We conclude that the investigated thermometers are not reliable as devices to measure radiant temperature, cannot be used to assess body core temperature during exercise, but may be used as a screening device, with 37.5°C as a threshold for fever in emergency care settings. © 2023 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

**□ Keywords—**thermometry; tympanic temperature; fever; arteria temporalis; thermometer; exercise; calibration

### Introduction

The human body's temperature is significantly influenced by ambient conditions, whereas the temperature of the body core (brain, heart, lungs, and intestines) remains relatively unaffected by environmental factors (1). Body core temperature determination in hospital emergency settings is essential to assess the presence of fever. Fever is

Declaration of generative AI and AI-assisted technologies in the writing process: During the preparation of this work, the author(s) used ChatGPT to improve language and readability. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

RECEIVED: 22 June 2023; FINAL SUBMISSION RECEIVED: 2 October 2023;  
 ACCEPTED: 22 October 2023

a complex physiologic response triggered by infectious or aseptic stimuli and defined as a body core temperature exceeding 38.0°C. When fever has been established, the source of the temperature increase can be diagnosed and treated accordingly.

Hospitals use a wide variety of measurement devices to determine the body core temperature. In academic hospitals in The Netherlands, the most common systems are based on infrared measurements in the ear canal or the forehead, because these systems are easy to use. However, the reliability and validity of those systems is challenged (2–8). This may result in any necessary treatments being delayed or not being started at all, with all its consequences. Also, we received indications from emergency personnel in hospitals that the readings of the infrared systems they used did not always comply with rectal measurements, used as a check. Therefore, we sought to test the accuracy and validity of temporal artery and tympanic measurements. The total measurement error is usually expressed as bias  $\pm$  2 SD, determining the 95% limits of agreement (LoA) between a measure and its reference. The acceptable LoA depend on the purpose and conditions of the measurement. In general, a 95% LoA of  $< \pm 0.5^\circ\text{C}$  is proposed as acceptable for spot checks like fever detection (9). Similarly, a good rule of thumb for the level of accuracy that is clinically necessary and that has been used in many studies, is that the inaccuracy of a thermometer should not exceed 0.5°C (10).

A previous study established that infrared tympanic measurement systems underestimate rectal temperature (T-RECT) and thus fail to detect patients with fever (T-RECT  $\geq 38^\circ\text{C}$ ) (11). They proposed to set the threshold for fever of the tympanic measurements to 37.5°C instead of 38.0°C to reduce the number of missed fevers, but obviously to the expense of more false positives. In our study, we also evaluated this threshold change and also determined if the underestimation might be due to inherent calibration errors in the infrared measurement systems.

Therefore, we decided to perform three experiments to evaluate the systems, ranging from basic accuracy to practical validity. To establish the accuracy, the three devices were calibrated using a certified infrared calibration system to determine if the devices measured the infrared radiation correctly. Second, we conducted a laboratory experiment in which we maintained a fixed T-RECT after cycling exercise and measured the body core temperature using the three infrared devices. This helped us determine the extent of deviation under controlled simulated fever conditions. Exercise is a less invasive way to change body core temperature than inflicting fever, and can be better controlled at fixed T-RECT values. Third, we performed measurements in patients at the emergency department (ED) that were suspected for infection after triage.

The first hypothesis was that the three devices would accurately measure the radiant temperature with a deviation of  $< 0.5^\circ\text{C}$  from the actual radiant temperature. The second hypothesis was that the three systems would underestimate the end-exercise rectal temperature by more than 0.5°C due to factors such as the ear thermometers capturing not only the tympanic membrane but also the surrounding ear canal wall, and sweating of the forehead reducing the temperature of the skin covering the temporal artery (12,13). The third hypothesis was that the three systems would underestimate the reference rectal temperature in emergency care settings by more than 0.5°C (11,13).

## Methods

### *Temperature Measurement Systems*

Three systems in use in academic hospitals in Groningen, Leiden, and Amsterdam were evaluated: the Braun Thermoscan® PRO 6000 (Welch Allyn, Skaneateles Falls, NY) and Genius 3 (Cardinal Health, Dublin, OH) aiming for the tympanic membrane, and Exergen TAT-5000S-RS232 (Exergen, Watertown, MA) aiming at the temporal artery.

### *Calibration*

Two samples of the Genius, Braun, and Exergen thermometers were calibrated using the Voltcraft IRS-350 (Voltcraft, Hirschau, Germany), a blackbody infrared calibrator with an emission coefficient of 0.95. This emission coefficient closely resembles the human skin and is independent of skin tone (14,15). The Genius has multiple measuring modes: rectal, ear, or oral. The ear mode is the unadjusted mode and was used during the calibration and during the protocol. For the calibration, the Genius and Braun were held in front of the calibrator, whereas the Exergen was slid from left to right in front of the calibrator. The thermometers were calibrated from 36°C to 41°C, with a 0.5°C interval. On every calibration point, the average of two measurements for each thermometer was noted.

### *Participants Exercise Study*

Nine men and three women participated in the study. The mean age of the participants was  $22 \pm 3$  years, the mean weight was  $77.1 \pm 14.8$  kg, and mean height was  $179 \pm 9$  cm. All participants signed an informed consent form. The study was approved by the local ethical committee (approval number VCWE-2021-075).

Exclusion criteria for this experiment were: age younger than 18 years, pregnancy, heart and vascular diseases, congenital or acquired anomaly of the external auditory meatus, and diseases of the ear. None of the patients used medication that might affect thermoregulation. Participants were instructed to wear sports clothing, that is, a t-shirt, shorts, and sneakers.

### *Procedures Exercise Study*

The experiment took place in the Human Performance Laboratory of Vrije Universiteit Amsterdam. Core temperatures were measured with a Mon-a-Therm 400TM temperature probe (rectal) (Medtronic, Minneapolis, MN), the Genius 3, the Braun ThermoScan PRO 6000, and the Exergen TAT-5000S. For each of the thermometers (with exception of the disposable rectal thermometer), two samples were made available and were used intermittently during the trials balanced across subjects.

The participant inserted the rectal thermometer approximately 15 cm beyond the anal sphincter. The rectal sensor was connected with a DT830B digital multimeter (All-Sun, Zhangzhou, China) to read the resistance, which was converted to temperature. The Genius 3 and the Braun ThermoScan PRO 6000 were used to measure the tympanic temperature in the left and right ear, whereas the Exergen TAT-5000S measured the temperature of the arteria temporalis. The corresponding user manuals were studied carefully in advance. The Exergen was moved over the skin from the center of the forehead to either hairline with the measure button pressed during the movement.

The participants cycled on a Lode Excalibur ergometer (Lode, Groningen, The Netherlands), starting with a 10-min warming-up phase at 1 W/kg body weight. The load increased every 2 min by 0.2 W/kg bodyweight, with cycling terminated 5 min after reaching a rectal temperature of 38.0°C. After warm-up, measurements were conducted each 2 min. Participants were free to choose the cadence. Participants were instructed not to wipe off sweat from their foreheads and were not allowed to drink during cycling.

### *Participants ED Study*

The study population consisted of 133 patients admitted to the ED of Amsterdam University Medical Centers between May and September 2022, whose temperature had to be measured in accordance with the applied triage protocol appropriate to the entry complaint or at the request of the attending physician. The mean age of the participants was  $61 \pm 17$  years, the mean weight was  $79.5$

$\pm 17.5$  kg, and mean height was  $173 \pm 11$  cm. Patients were not checked for medication.

We used G\*Power version 3.1 to calculate the required number of test subjects. The manuals of the investigated systems reported an accuracy of  $\pm 0.2^\circ\text{C}$  for Braun (Braun ThermoScan® PRO6000, Product/sales literature 901083, 2005; Kaz USA Inc., Marlborough, MA) and  $\pm 0.1^\circ\text{C}$  for Exergen (Exergen Document p/n 818641r5, 2021). The accuracy of the Genius is reported as  $\pm 0.1^\circ\text{C}$  (Cardinal Health User Manual 03/2018 – H9950 – WF#776288). Based on these values, a desired accuracy of  $0.12^\circ\text{C}$  was chosen for the statistical power analysis. Calculations were made with a SD of  $0.49^\circ\text{C}$ , based on a small trial set. This gave an effect size (Cohen's d) of 0.245. Calculations were made with an alpha of 0.05 and a power of 0.8. This showed that the number of participants required was 133.

Participants had to meet the following criteria to be included in this study: participants had to be 18 years or older; measurement of temperature had to be necessary according to the applied triage protocol, it was required to have proficiency in Dutch or English language, or to be accompanied by an (in)formal caregiver proficient in these languages. Participants who met any of the following criteria were excluded: inability to give consent due to life-threatening situations or reduced consciousness; incapacitated patients (i.e., patients who did not understand the information about their illness, could not decide for themselves, or did not understand the consequences of a decision); pregnancy; congenital or acquired anomaly of the external auditory meatus diseases of the ear; participants with hemorrhoids to avoid discomfort and possible bleeding during rectal measurements. The study was approved by the Medical Ethical Commission of Amsterdam University Medical Centers under number NL79962.018.22.

### *Procedures ED Study*

Prior to the measurements, the student researchers performed preparatory work for 5 consecutive days at Vrije Universiteit Amsterdam to improve interobserver agreement. Besides, these practicing days were necessary to become familiar with the testing methods, avoiding any inconvenience during the measurements in the hospital. Five days of training improved interobserver variability from  $0.19^\circ\text{C}$  to  $0.14^\circ\text{C}$ .

The participants were asked to estimate their height and weight. Previous research with 562 Dutch men and 695 Dutch women, representative of the Dutch population, showed that the self-reports correspond well with actual height and weight, indicating that our research population is an unbiased representation of the Dutch population (16).

Upon a patient's arrival at the ED, specialized nursing staff conducted triage. If the triage protocol or the request of the emergency physician necessitated a temperature measurement for further treatment determination, the triage nurse explained the ongoing thermometry validity research to the patient. The nurse then sought the patient's consent for participation in the study. After obtaining consent, a student researcher approached the patient and asked if they were able to perform the rectal measurement themselves using the ThermoVal® Rapid thermometer (Hartmann, Heidenheim, Germany). If this was not possible, a nurse performed the rectal measurement. The tip of the thermometer was inserted 2 cm into the rectum. This measurement took about 10–15 s. The other six temperature measurements (Braun, Genius, and Exergen, left and right) were performed in random order in accordance with the respective manuals. Disposable caps for all thermometers were replaced after every patient. Ambient temperature was not measured, but estimated at about 20°C.

### Data Processing

The data were analyzed using Statistica (17). Normality of the data was assessed using the Kolmogorov-Smirnov test. Left-right differences per infrared device (Braun, Genius, Exergen) and the difference between rectal temperature and each infrared device were analyzed using *t*-tests. In cases where no difference between left and right was observed, the data were pooled. Significance level was set at  $p < 0.05$ .

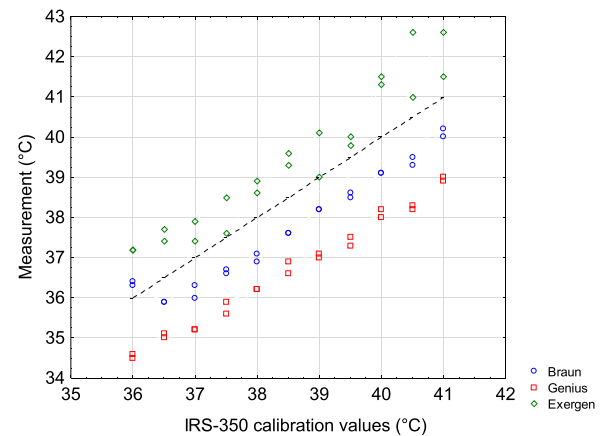
## Results

### Calibration

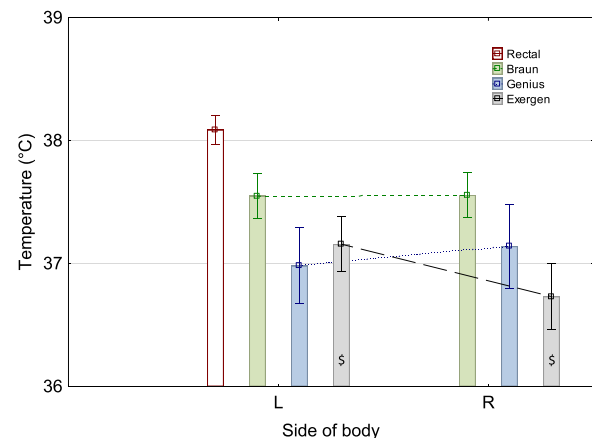
Figure 1 shows the results of the calibration. Braun and Genius underestimated radiant temperature by  $0.77 \pm 0.39^\circ\text{C}$  and  $1.84 \pm 0.26^\circ\text{C}$ , respectively, and Exergen overestimated radiant temperature ( $0.90 \pm 0.51^\circ\text{C}$ ) ( $p < 0.001$ , *t*-test for single means).

### Exercise Study

The mean temperatures of all systems at the end of exercise are shown in Figure 2. All temperatures were normally distributed ( $p > 0.05$ ). Rectal temperature averaged  $38.1 \pm 0.2^\circ\text{C}$ . Braun temperatures did not differ between left and right, values were  $37.6 \pm 0.3^\circ\text{C}$  for both sides. Genius temperatures also did not differ between left and right with  $37.0 \pm 0.5^\circ\text{C}$  and  $37.1 \pm 0.5^\circ\text{C}$ , respectively. The Exergen temperature was significantly higher on the left than the right side ( $p = 0.01$ ) with values of



**Figure 1. Calibration results: values from Voltcraft IRS-350 system on horizontal axis and values of two Braun Thermoscan® PRO 6000, Genius 3, and Exergen TAT-5000S-RS232 systems on vertical axis.**



**Figure 2. Mean temperature and 95% confidence interval end exercise for rectal temperature, Braun Thermoscan® PRO 6000, Genius 3, and Exergen TAT-5000S-RS232 systems for left (L) and right (R) side of body.  $n = 12$ . \$ = different between L and R.**

$37.2 \pm 0.4^\circ\text{C}$  and  $36.7 \pm 0.4^\circ\text{C}$ , respectively. Differences between rectal temperature and the temperature of each of the three investigated systems differed significantly from 0 ( $p < 0.001$ , *t*-test for single means).

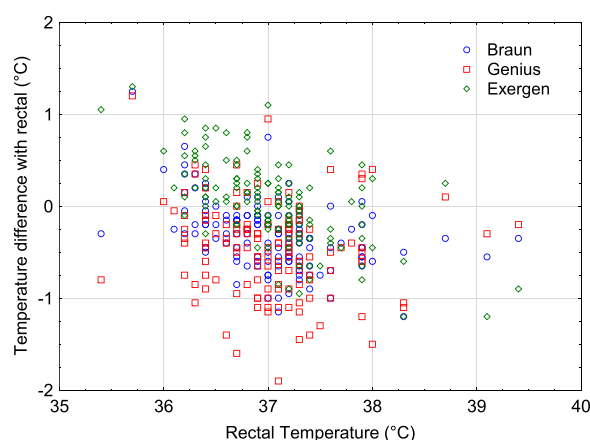
### ED Study

Braun, Genius, and Exergen values did not significantly differ between left and right ( $p > 0.05$ ) and were therefore combined in the analysis. All temperatures were normally distributed ( $p > 0.05$ ). Rectal temperature had a range of  $35.4^\circ\text{C}$  to  $39.4^\circ\text{C}$  (mean  $37.03 \pm 0.60^\circ\text{C}$ ). The overall average difference from rectal temperature was  $-0.31 \pm 0.37^\circ\text{C}$  (Braun),  $-0.46 \pm 0.50^\circ\text{C}$  (Genius), and  $0.04 \pm 0.46^\circ\text{C}$  (Exergen). Seven patients had a rectal temperature  $\geq 38^\circ\text{C}$  and were classified as fever patients.

**Table 1. Mean Temperatures ( $\pm$  Standard Deviation) of Rectal Temperature, Braun Thermoscan® PRO 6000, Genius 3, and Exergen TAT-5000S-RS232 Systems for Patients Without Fever (2<sup>nd</sup> Column) and with Fever (3<sup>rd</sup> Column), and Differences Between Rectal Temperature and Braun, Genius, and Exergen Temperature for Patients Without Fever (4<sup>th</sup> Column) and with Fever (5<sup>th</sup> Column).**

	No Fever (n = 126)	No Fever $\Delta$ Rectal (n = 126)	Fever (n = 7)	Fever $\Delta$ Rectal (n = 7)
Rectal	36.95 $\pm$ 0.48		38.54 $\pm$ 0.54	
Braun	36.65 $\pm$ 0.45*	−0.30 $\pm$ 0.37*	38.02 $\pm$ 0.68	−0.52 $\pm$ 0.34
Genius	36.49 $\pm$ 0.61*	−0.46 $\pm$ 0.49*	38.02 $\pm$ 1.03	−0.52 $\pm$ 0.70
Exergen	37.02 $\pm$ 0.41	0.07 $\pm$ 0.43	38.00 $\pm$ 0.63	−0.54 $\pm$ 0.62

\* Indicates significant differences from rectal temperature.



**Figure 3. Braun Thermoscan® PRO 6000, Genius 3, and Exergen TAT-5000S-RS232 temperature values minus rectal temperature plotted against rectal temperature.**

Figure 3 and Table 1 show the measurement data for the fever and nonfever group separately.

The ear thermometers significantly underestimated rectal temperature for the nonfever group ( $p < 0.001$ ). The Exergen values were not different from rectal temperature measurements in this group ( $p > 0.05$ ). The seven subjects with fever (rectal temperature  $\geq 38^\circ\text{C}$ ) had values that were  $0.52 \pm 0.34^\circ\text{C}$  (Braun),  $0.52 \pm 0.70^\circ\text{C}$  (Genius), and  $0.54 \pm 0.62^\circ\text{C}$  (Exergen) lower than rectal temperature. There was no significant difference from rectal temperature ( $p > 0.05$ ).

## Discussion

The first hypothesis was that the three devices would accurately measure radiant temperature with a deviation of  $< 0.5^\circ\text{C}$  from the actual radiant temperature. This hypothesis has to be rejected; values were systematically higher (Exergen) or lower (Braun and Genius) than the values of

the Voltcraft IRS-350 calibration system. Ambient conditions during calibration were in line with International Organization for Standardization standard IEC 80601-2-59 within the ambient temperature range of  $18^\circ\text{C}$  to  $24^\circ\text{C}$  and humidity range of 10% to 75% (18). The Voltcraft IRS-350 calibration source has a display step size of  $0.1^\circ\text{C}$ , a temperature range of  $50$ – $350^\circ\text{C}$ , an accuracy of  $\pm 0.5^\circ\text{C}$  at  $100^\circ\text{C}$ , and a stability of  $\pm 0.1^\circ\text{C}$  at  $100^\circ\text{C}$ . Even though our temperature settings of  $36$ – $41^\circ\text{C}$  were outside the specified range of  $50$ – $350^\circ\text{C}$ , the values exactly matched the values read by the FLIR TC-167 camera (FLIR Systems Inc., Wilsonville, OR) with a resolution of  $0.1^\circ\text{C}$  and accuracy of  $\pm 1.5^\circ\text{C}$ . Also, the temperature was stable within  $0.1^\circ\text{C}$  for the black radiator surface area. We concluded that the investigated three systems do not accurately measure radiation temperature. The standards related to assess body core temperature mainly refer to measurement of the inner canthus of the eye for body core temperature assessment, a method that has been shown not to be accurate (18–21). ASTM standards (ASTM International, West Conshohocken, PA) related to infrared thermometry are applicable in the United States, but not in Europe (22,23).

The second hypothesis was that the three systems would underestimate the rectal temperature after exercise by more than  $0.5^\circ\text{C}$ . This hypothesis is confirmed. All systems showed a significant underestimation ( $0.54 \pm 0.27^\circ\text{C}$  Braun,  $1.03 \pm 0.48^\circ\text{C}$  Genius,  $1.14 \pm 0.38^\circ\text{C}$  Exergen). It was remarkable that the Exergen values differed between the left and right side of the forehead ( $37.2 \pm 0.4^\circ\text{C}$  and  $36.7 \pm 0.4^\circ\text{C}$ , respectively). A possible explanation may be that the measurer was standing at the right side of the body and may have moved the measuring device differently over the left and right forehead. From an anatomical point of view, no differences were expected between the left and right temporal artery. The underestimation of infrared ear thermometers is reported prior and



may be related to the fact that the device does not only see the tympanic membrane, but also the surrounding ear canal wall that is generally colder (12). The underestimation of the Exergen system may be related to evaporating sweat at the forehead that reduces the temperature of the skin covering the temporal artery (13). Thermal capsules for exercise body core temperature assessment are a good alternative for rectal temperature in operational settings (24–27). Heat flux systems for core temperature assessment during exercise have been shown to have an offset with rectal temperature, but may be used after correction (28).

The third hypothesis was that the three systems would underestimate rectal temperature in emergency care settings by more than 0.5°C. This hypothesis was confirmed for the ear thermometers, but not for Exergen. For the 7 patients with fever, all systems showed about 0.5°C lower mean temperatures than rectal, with considerable standard deviations ranging from 0.34°C (Braun) to 0.70°C (Genius). Based on similar findings, Mogensen et al. suggested setting the threshold for fever 0.5°C lower for these devices (11). When we performed the same shift for our systems, the results show sensitivity ranging from 57% to 86% and high specificity of 90–97% (Table 2) (11). It was previously discussed that tests can be considered useful when the sum of sensitivity and specificity exceeds 150%, which was the case in our study for all investigated systems (29). We performed an analysis for thresholds of 37.1°C to 38.0°C in steps of 0.1°C that showed that the sum of sensitivity and specificity peak at 166% for thresholds of 37.1°C to 37.5°C. The highest value was for the Braun system, with 182% for a threshold of 37.5°C.

In clinical settings there is a strong need for fast and accurate body core temperature assessment. Life-threatening bacterial infections are accompanied by strong increases in body core temperature and should be treated accordingly. Although the investigated infrared devices have been shown not to meet the  $\pm 0.5^\circ\text{C}$  threshold for (clinical) accuracy, they may be used in screening protocols. The high specificity that we observed in our study means that application of the revised threshold for fever of 37.5°C using the three devices, leads to the elimination of the major part of the population without fever. However, because the sensitivity of the investigated thermometers may be as low as 57%, it may be considered to lower the threshold to 37.3°C, realizing that the specificity will then be reduced. To overcome low specificity, we recommend obtaining a rectal temperature if the threshold of 37.3°C is reached with one of these three thermometers.

### Limitations

The number of subjects with fever in our ED study was low, but the findings are in general agreement with the pre-

**Table 2. Overview of Sensitivity and Specificity of Systems Used by Mogensen et al. [2018; (11)] and the Braun Thermoscan® PRO 6000, Genius 3, and Exergen TAT-5000S-RS232 Systems Evaluated in this Study\*.**

Test ( $\geq 37.5^\circ\text{C}$ )	Morgensen et al. (2018)		Braun		Genius		Exergen	
	Fever (T-RECT $\geq 38.0^\circ\text{C}$ )		Fever (T-RECT $\geq 38.0^\circ\text{C}$ )		Fever (T-RECT $\geq 38.0^\circ\text{C}$ )		Fever (T-RECT $\geq 38.0^\circ\text{C}$ )	
	Present	Absent	Present	Absent	Present	Absent	Present	Absent
Positive	39	94	5	5	4	4	6	13
Negative	2	505	2	121	3	122	1	113
Sensitivity (%)		95		71		57		86
Specificity (%)		84		96		97		90

\* Patients were classified as fever patients when rectal temperature (T-RECT)  $\geq 38^\circ\text{C}$ . The test was when systems indicated a temperature  $> 37.5^\circ\text{C}$ . For instance, the Braun system detected 5 of 7 patients with fever (sensitivity 71%) and detected 121 of 126 patients without fever (specificity 96%). T-RECT = rectal end-exercise temperature.

vious study in which 41 patients with fever were included (11). The lower sensitivity in our study may be related to the small number of subjects in our study with fever.

Ambient temperature was not measured in the ED during the experiment, but measurements later confirmed that the temperatures were close to 20°C.

## Conclusions

We conclude that the investigated thermometers are not reliable as devices to measure radiant temperature; they cannot be used to assess body core temperature during exercise, but may be used as a screening device with 37.5°C as a threshold for fever in emergency care settings.

## Declaration of competing interest

The authors of this study have no conflicts of interest to report.

## Acknowledgements

We would like to acknowledge the staff of the Emergency Department of the Amsterdam UMC location AMC for their cooperation, and a special thanks to Ruud Voorn, technician, and Dr. Milan Ridderikhof PhD and MD Emergency Medicine, for their contribution to this study.

## References

1. Periard JD, Eijssvogels TMH, Daanen HAM. Exercise under heat stress: thermoregulation, hydration, performance implications, and mitigation strategies. *Physiol Rev* 2021;101:1873–979.
2. Cutuli SL, See EJ, Osawa EA, et al. Accuracy of non-invasive body temperature measurement methods in adult patients admitted to the intensive care unit: a systematic review and meta-analysis. *Crit Care Resusc* 2021;23:6–13.
3. 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.
4. McCarthy P, Heusch A. The vagaries of ear temperature assessment. *J Med Eng Technol* 2006;30:242–51.
5. Niven DJ, Gaudet JE, Laupland KB, Mrklas KJ, Roberts DJ, Stelfox HT. Accuracy of peripheral thermometers for estimating temperature: a systematic review and meta-analysis. *Ann Intern Med* 2015;163:768–77.
6. Pusnik I, Drnovsek J. Infrared ear thermometers—parameters influencing their reading and accuracy. *Physiol Meas* 2005;26:1075–1084.
7. Roth RN, Verdile VP, Grollman LJ, Stone DA. Agreement between rectal and tympanic membrane temperatures in marathon runners. *Ann Emerg Med* 1996;28:414–17.
8. Stavem K, Saxholm H, Smith-Erichsen N. Accuracy of infrared ear thermometry in adult patients. *Intensive Care Med* 1997;23:100–5.
9. Wartzek T, Muhlsteff J, Imhoff M. Temperature measurement. *Biomed Tech (Berl)* 2011;56:241–57.
10. Sessler DI. Temperature monitoring and perioperative thermoregulation. *Anesthesiology* 2008;109:318–38.
11. Mogensen CB, Vilhelmsen MB, Jepsen J, Boye LK, Persson MH, Skyum F. Ear measurement of temperature is only useful for screening for fever in an adult emergency department. *BMC Emerg Med* 2018;18:51.
12. Daanen HAM. Infrared tympanic temperature and ear canal morphology. *J Med Eng Technol* 2006;30:224–34.
13. Kistemaker JA, Den Hartog EA, Daanen HA. Reliability of an infrared forehead skin thermometer for core temperature measurements. *J Med Eng Technol* 2006;30:252–61.
14. Steketee J. Spectral emissivity of skin and pericardium. *Phys Med Biol* 1973;18:686–94.
15. Simpson CR, Kohl M, Essenpreis M, Cope M. Near-infrared optical properties of ex vivo human skin and subcutaneous tissues measured using the Monte Carlo inversion technique. *Phys Med Biol* 1998;43:2465–78.
16. Krul AJ, Daanen HAM, Choi H. Self-reported and measured weight, height and body mass index (BMI) in Italy, the Netherlands and North America. *Eur J Public Health* 2011;21:414–19.
17. TIBCO Software Inc. (2020). Data Science Workbench, version 14. <http://tibco.com>. Accessed December 12, 2023.
18. International Organization for Standardization (ISO). ISO IEC 80601-2-59. IEC 80601-2-59- Medical electrical equipment – Part 2-59: Particular requirements for the basic safety and essential performance of screening thermographs for human febrile temperature screening. Geneva: ISO; 2017.
19. International Organization for Standardization (ISO). ISO TR 13154. ISO/TR 13154:2017 Medical electrical equipment — Deployment, implementation and operational guidelines for identifying febrile humans using a screening thermograph. Geneva: ISO; 2017.
20. Mekjavic IB, Tipton MJ. Myths and methodologies: degrees of freedom – limitations of infrared thermographic screening for Covid-19 and other infections. *Exp Physiol* 2022;107:733–42.
21. Teunissen LP, Daanen HAM. Infrared thermal imaging of the inner canthus of the eye as an estimator of body core temperature. *J Med Eng Technol* 2011;35:134–8.
22. ASTM E1965-98. Standard specification for infrared thermometers for intermittent determination of patient temperature. West Conshohocken, PA: ASTM; 2016.
23. ASTM E2847-13. Standard practice for calibration and accuracy verification of wideband infrared thermometers. West Conshohocken, PA: ASTM; 2021.
24. Bongers CCWG, Daanen HAM, Bogerd CP, Hopman MTE, Eijssvogels TMH. Validity, reliability, and inertia of four different temperature capsule systems. *Med Sci Sports Exerc* 2018;50:169–75.
25. Byrne C, Lim CL. The ingestible telemetric body core temperature sensor: a review of validity and exercise applications. *Br J Sports Med* 2007;41:126–33.
26. Gant N, Atkinson G, Williams C. The validity and reliability of intestinal temperature during intermittent running. *Med Sci Sports Exerc* 2006;38:1926–31.
27. Kolka MA, Quigley MD, Blanchard LA, Toyota DA, Stephenson LA. Validation of a temperature telemetry system during moderate and strenuous exercise. *J Therm Biol* 1993;18:203–10.
28. Daanen HAM, Kohlen V, Teunissen LPJ. Heat flux systems for body core temperature assessment during exercise. *J Therm Biol* 2023;112.
29. Power M, Fell G, Wright M. Principles for high-quality, high-value testing. *Evid Based Med* 2013;18:5–10.