ORIGINAL ARTICLE

Performance of infrared ear and forehead thermometers: a comparative study in 205 febrile and afebrile children

Patricia A Hamilton, Lorenzo S Marcos and Michelle Secic

Aims and objectives. This study compared readings from two professional-grade, commercially available infrared (IR) thermometers, the ThermoScan\textsuperscript{R} PRO 4000 prewarmed tip ear thermometer and the Temporal Scanner\textsuperscript{TM} TAT-5000 temporal artery thermometer.

Background. The repeatability and precision of readings from IR thermometers for professional use were questioned in the past, but in recent years, these types of thermometers have been technologically improved, so their ability to replicate standard temperature readings reliably should be re-examined.

Design. Febrile and afebrile children were recruited from the emergency department, overflow treatment areas and the paediatric intensive care unit of a large hospital in Argentina. Each child had a randomised sequence of seven temperature readings, including three from the ear, three from the forehead or behind the ear and one reference oral or rectal reading.

Methods. Temperature readings were taken with the ThermoScan PRO 4000, the Temporal Scanner TAT-5000 and the monitor mode of SureTemp\textsuperscript{R} Plus, a widely used professional-grade contact thermometer, for reference.

Results. Of 205 children, 46\% were febrile, per reference thermometer readings. While mean ThermoScan PRO 4000 febrile measurements did not differ significantly from reference, mean Temporal Scanner TAT-5000 febrile measurements were significantly lower (by a mean of 0.42 °C) than the reference. Overall bias of the ThermoScan PRO 4000 was significantly lower than that of the Temporal Scanner TAT-5000; repeatability was 1.5 times higher, and overall false-negative rate was about a third that of Temporal Scanner TAT-5000, when compared to the reference.

Conclusions. This study indicates that the ThermoScan PRO 4000 provides measurements closer to those of a professional-grade contact thermometer when compared to the Temporal Scanner TAT-5000.

Relevance to clinical practice. The odds of misclassifying a febrile child as non-febrile were about threefold higher with the Temporal Scanner TAT-5000 temporal artery thermometer than with the ThermoScan\textsuperscript{R} PRO 4000 ear thermometer.

Key words: ear thermometer, fever, forehead thermometer, infrared thermometry, oral thermometer, temporal artery thermometer

Accepted for publication: 28 August 2012

Introduction

Various types of infrared (IR) fever thermometers are now available for rapid, non-invasive temperature measurement in the doctor’s office, clinic or hospital. This study was designed to assess the bias and precision of two professional-grade IR technologies (an ear thermometer and a temporal artery thermometer) compared to a traditional widely used thermistor-based contact thermometer of the type most clinicians are familiar with.

Authors: Patricia A Hamilton, RN, MS, President, ClinStat Consulting LLC, Cardiff by the Sea, CA, USA; Lorenzo S Marcos, MD, Principal Investigator, Jefe de Trabajos Prácticos, Hospital del Niño Jesús, San Miguel de Tucuman, Argentina; Michelle Secic, MS, President, Secic Statistical Consulting, Inc., Chardon, OH, USA

Correspondence: Patricia A Hamilton, President, ClinStat Consulting LLC, 2492 Manchester Avenue, Cardiff by the Sea, CA 92007, USA. Telephone: +1 760 207 5260. E-mail: pb@clinstactconsulting.com

© 2013 Blackwell Publishing Ltd
Background

Infrared ear thermometers, which were introduced to the clinical market in 1986, initially had moderate acceptance by professionals because of concerns that they may not give the same readings as thermometers in traditional measurement sites (oral, rectal and axillary). (Chamberlain et al. 1991, Muma et al. 1992, Bernardo et al. 1999, Craig et al. 2002, Nimah et al. 2006). Disparities in single-point-in-time temperature measurements from different sites may be attributable to factors such as probe placement, mouth breathing, drinking and smoking (which affect oral readings) (Tandberg & Sklar 1983, Terndrup et al. 1989); ambient temperature near the skin surface (which affects axillary and forehead readings) (Erickson & Kirklin 1993, Cusson et al. 1997, Crawford et al. 2005); and rate of temperature change at different body sites (which may differ from rate of temperature change in core body temperature) (Ash et al. 1992, Twerenbold et al. 2010).

During the past two decades, ear thermometer technology has improved. The prewarmed tip IRHTET used for this study has a heating element in the sensor that warms the probe tip to just under normal body temperature in order to prevent the probe from cooling the ear canal and adjacent tissue during measurements, and the symmetric sensor is designed to reduce effects of temperature gradients. The thermometer is inserted into the outer cartilaginous part of the auditory canal; it does not touch the osseous part or the tympanic membrane. The IR sensor is located closer to the tip of the probe than in earlier models, thereby facilitating the activation of the prewarmed sensor tip. A positive feedback system in the IRHTET informs the user of errors in positioning of the thermometer in the ear.

Another contemporary technology for rapid, non-invasive temperature measurement is temporal artery thermometry. The forehead over the temporal artery is easily accessible for measurements. The temporal artery thermometer (TAT) used in this study assesses temperature using IR technology to detect the heat naturally emitted from the skin surface of the forehead over the temporal artery. The TAT incorporates a patented arterial heat balance system that automatically accounts for the effects of ambient temperature on the skin.

In this clinical study, the performance of the IRHTET was compared to that of the TAT in febrile and afebrile children. Measurements from both thermometers were compared with oral and rectal measurements taken in monitor mode with a reference thermometer – a traditional, commercially available, professional contact thermometer that uses a high-quality thermistor sensor in the probe. The aim was to determine whether measurements using the ear thermometer might more closely reflect reference thermometer readings than temporal artery measurements. Because mean body temperature (Herzog & Phillips 2011) and degree of cooperation are affected by age, the influence of age on readings with each of the thermometers was also assessed. As expected from the literature (Herzog & Phillips 2011), mean temperatures were somewhat higher for children aged two to four years than for older children.

Methods

Subjects and thermometers

The thermometers investigated in this study were the ThermoScan® PRO 4000 IR prewarmed tip ear thermometer, designated as IRHTET (Braun GmbH, Kronberg, Germany), and the Temporal Scanner™ TAT-5000 IR temporal artery thermometer, designated as TAT (Exergen Corp, Watertown, MA, USA). Both are IR thermometers that are used in predictive mode and can provide a measurement in seconds using an algorithm that extrapolates from the speed at which temperature changes as the thermometer warms up. In contrast to these thermometers, the SureTemp® Plus electronic contact thermometer (Welch Allyn, Skaneateles Falls, NY, USA) was used in monitor mode and was designated as the reference thermometer. The SureTemp® Plus (Welch Allyn) in monitor mode was chosen as the reference thermometer for several reasons: In a study comparing various thermometer types to core temperature readings from a Foley catheter with a thermistor, electronic oral thermometry was the most accurate and reliable, with mean ± SD difference of -0.25 ± 0.38 °C (Langham et al. 2009). When oral measurements were taken by trained ICU nurses in a standardised fashion, the SureTemp® Plus was shown to provide measurements that differed by an average of <0.1 °C from the pulmonary artery temperature (Lawson et al. 2007). Taking a three-minute reading with SureTemp® Plus in monitor mode ensures that any bias that might be introduced by an algorithm will not influence the reading, which is a direct reflection of measurement by a calibrated thermistor. The SureTemp® Plus has been used as a reference standard in other thermometer comparisons as well (Braun et al. 1998, Giuliano et al. 2000, Mangat et al. 2010).

Prior to the clinical study, all thermometers were calibrated by their respective manufacturers in individual laboratory settings to assure reliability and laboratory accuracy. Prior to subject enrolment each day, ear thermometers and forehead thermometers were verified in their respective
calibration devices to assure that the thermometers were functioning properly and measuring within specifications. The reference thermometers (monitor mode) were checked for proper calibration prior to study initiation.

The charge nurse (or equivalent) recruited children from the emergency department and the outpatient patient treatment areas of the Hospital del Niño Jesus in Tucuman, Argentina, during March of 2007. Screening for inclusion did not include temperature measurements. To ensure that sufficient numbers of febrile participants were recruited, hospital nurses recruited additional febrile children from inpatients in the paediatric intensive care unit. The recruiting nurse ascertained from parents or guardians whether they were interested in participating. Approval for this study was granted by the Human Subjects Protection Review Board of the Hospital del Niño Jesus prior to subject enrolment. Written informed consent was obtained from the parent or guardian, and verbal assent, when appropriate, was obtained from the child prior to enrolment. Children were enrolled without regard to gender or ethnic background.

For purposes of this study, children were grouped into four age groups: (1) 0–24 months, (2) >24–48 months, (3) >48 months–10 years and (4) >10–18 years. Participants were accepted into the study until an age category was filled. Half of the study population was to be four years old and younger, as young children have problems communicating their symptoms, and their temperature measurements can change rapidly.

Children with obvious deformities of the face or ear canal were excluded. Subjects were not screened for otitis media, cerumen or other conditions of the ear and ear canal, as studies have shown that the presence of otitis media or cerumen affect ear temperature measurements only slightly (Brennan et al. 1994, Twerenbold et al. 2010).

The definition of fever for this study was a temperature ≥ 38.0 °C (100.4°F) (European Committee for Standardization 2003) measured by the reference thermometer in monitor mode for at least three minutes. Prior to taking forehead measurements, the area for taking measurements was prepared in accordance with the manufacturer’s instructions. Each enrolled participant had a temperature measurement series consisting of seven temperature measurements: three in the ear, three on the forehead or behind the ear and a three-minute (monitor mode) oral or rectal measurement as a control. The same ear and forehead thermometers were used for triplicate measurements. The order of use for both the ear and the forehead thermometer was randomised. The same reference thermometer (using the appropriate oral or rectal probe) was used for all oral and rectal measurements.

Children five years of age and under had a three-minute rectal temperature measurement, and children greater than five years of age had a three-minute oral temperature control measurement for comparison. A successful temperature measurement series was defined as obtaining all seven temperature measurements. Inability to obtain all seven measurements was considered cause for exclusion of the data.

All measurements were obtained in the presence of a physician observer by the same right-handed registered nurse, as recommended by the European Committee for Standardization (2003). The nurse was well trained on all equipment and had years of experience taking thousands of temperature measurements. All measurements (unless contraindicated) were measured on the right forehead/temple/behind the ear area and in the right ear. There was at least a one-minute wait time between temperature measurements, but never more than a five-minute period between triplicate measurements. During the wait period between measurements, the thermometers were placed on a bedside table to prevent any possible effect of hand-warming on the devices.

During all oral, ear and forehead measurements, the nurse stabilised the subject’s head with her left hand to prevent the child from pulling away during the measurement, which helped assure more accurate measurements. All equipment was thoroughly cleaned according to the manufacturers’ instructions at the end of each work day. Appropriate manufacturer probe covers were used and were changed for each measurement for each subject. Welch Allyn SureTemp oral probes with probe covers and rectal probes with probe covers were used for all oral and rectal measurements. Probe covers manufactured by Braun were used on the IRHTET. The TAT thermometer was cleaned with an alcohol wipe after triplicate measurements were completed on each subject.

Statistical methods

The primary end point of the study was to compare the bias (with corresponding standard deviation) relative to the reference thermometer for the TAT vs. the IRHTET, as recommended by the European Committee for Standardization (2003). A secondary end point was to compare the TAT vs. the IRHTET regarding the likelihood of false-negative and false-positive fever classification. A post hoc analysis of the repeatability of both instruments was also carried out.

Descriptive statistics were summarised as means and standard deviations for continuous variables and as frequencies and per cent for categorical variables; 95% confidence intervals as well as graphical summaries were also
used to describe the data. Results were considered statistically significant if the $p$-value was less than or equal to an alpha level of 0.05. The required sample size for a difference of 0.2°C and a standard deviation of 0.6°C was calculated as 155. Subject enrolment numbers allowed for 10–20% attrition rate.

Bland–Altman plots were used to compare agreement between the two devices. To assess the relative performance differences of the IRHTET and the TAT, a fixed, two-factor repeated measures general linear model (GLM) was fit. Fixed factors in the model were age group and febrile status, as defined by $\geq 38.0^\circ C$ ($\geq 100.4^\circ F$) using the reference thermometer (as defined by the European Committee for Standardization 2003). Repeated measures were calculated as the difference of the average for each device less the reference thermometer measurement by subject. A negative value indicated the reference thermometer measurement was greater than the reading with the comparative device. Triplicate measures for the IRHTET and the TAT were averaged by subject. Within-subject averages were then subtracted from the reference measurement. If the two instruments were equivalent, the mean differences within and between subjects would not be measurably different.

To assess which device better classified febrile status relative to the reference thermometer measurement, McNemar’s test for symmetry was performed on the tabulation of febrile classification for each device compared to the reference thermometer.

Repeatability of triplicate measurements for the IRHTET and the TAT was assessed within age groups to determine whether the estimates were within 0–30°C (0–54°F), which is the standard of clinical repeatability set by the European Committee for Standardization (2003).

Results

Population summary

Two hundred and twelve children aged 0–18 years were enrolled, and 205 completed the study. Seven participants were excluded from the final analyses, as follows: gender was not identified for one enrolled participant; one participant had a colostomy; and three participants refused rectal measurements. Axillary temperature measurement is the standard at Hospital del Niño Jesús as well as the standard for the country in general; therefore, children were not accustomed to oral or rectal temperature measurements.) Two participants in the afebrile groups who were enrolled after their age category was filled were not included in the analysis. The overall population contained 58% boys. Forty-six per cent of the population was febrile ($\geq 38.0^\circ C$, $\geq 100.4^\circ F$) according to the reference thermometer measurement. The distribution of participants and gender appeared comparable among age groups, except that the older age group contained 70% males. Per cent febrile participants decreased somewhat as age increased (Table 1).

Bland–Altman analyses

Bland–Altman analyses (Bland & Altman 1986), classed by age group, were performed for (1) the IRHTET compared to the TAT (Table 2); (2) the IRHTET compared to the reference thermometer (Table 2, Fig. 1a); and (3) the TAT compared to the reference thermometer (Table 2, Fig. 1b). In all cases, the margin of error (half-width) was narrow, with larger margins of error consistently occurring in the comparison between the TAT and the reference thermometer (Table 2). Within comparisons to the reference thermometer readings, the average bias (difference) for overall comparisons ranged from 0–0.3°C (0–0.5°F), indicating an acceptable level of bias between instruments.

Instrument comparison

For both IRHTET and the TAT, triplicate averages correlated highly with values from the reference thermometer. The correlation coefficient for the IRHTET and reference thermometer was 0.925 ($p < 0.001$). The correlation coefficient for the TAT and reference thermometer was 0.828 ($p < 0.001$). The correlation coefficient for the IRHTET and the TAT was 0.879 ($p < 0.001$).

Results indicated that the overall instrument effect was significantly different, with the IRHTET having a mean difference of $-0.001^\circ C$ ($-0.002^\circ F$) from the reference thermometer and the TAT having a mean difference of $-0.17^\circ C$ ($-0.30^\circ F$) ($F_{1.197} = 30.1$, $p < 0.001$) from the reference thermometer.

Table 1 Population summary

<table>
<thead>
<tr>
<th>Age range</th>
<th>Subjects</th>
<th>Avg. age (month)</th>
<th>% male</th>
<th>% febrile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>205</td>
<td>66</td>
<td>58</td>
<td>46</td>
</tr>
<tr>
<td>0–24 months</td>
<td>56</td>
<td>13</td>
<td>54</td>
<td>54</td>
</tr>
<tr>
<td>&gt;24–48 months</td>
<td>50</td>
<td>36</td>
<td>52</td>
<td>44</td>
</tr>
<tr>
<td>&gt;4–10 years</td>
<td>56</td>
<td>81</td>
<td>59</td>
<td>46</td>
</tr>
<tr>
<td>&gt;10–18 years</td>
<td>43</td>
<td>153</td>
<td>70</td>
<td>40</td>
</tr>
</tbody>
</table>

Study population was deliberately concentrated in younger age groups.

© 2013 Blackwell Publishing Ltd

Journal of Clinical Nursing
thermometer. This suggests that, on average, the IRHTET measurements were closer to the reference thermometer measurements than those of the TAT. The standard deviation of the IRHTET was 0.39 °C (0.70 °F), while the standard deviation of the TAT was 0.58 °C (1.04 °F) when compared to the reference thermometer. Accordingly, the 95% confidence interval was smaller for the IRHTET compared to the TAT.

The interaction between age groups and instrument was not significantly different between thermometers. No trend up or down within or between thermometers was apparent in any age group, suggesting that no systematic bias was present with age comparing instruments ($F_{3,197} = 0.56$, $p = 0.64$) (Fig. 2).

The interaction of thermometer with febrile status was significant ($F_{1,197} = 25.1$, $p < 0.001$), indicating that the mean difference within age groups for febrile patients differed significantly between thermometers (Fig. 3). While mean (IRHTET) febrile measurements did not differ significantly (−0.08 °C) from reference, mean TAT febrile measurements were significantly lower (by 0.42 °C) than the reference.

The interaction of thermometer, age group and febrile status was not significantly different ($F_{3,197} = 0.7$, $p = 0.56$) between thermometers, indicating that the bias observed between instruments was consistent in all the age groups. While the instruments were significantly different within the febrile population, no systematic bias across age groups was seen in the statistical analysis.

Classification analysis

Using the reference thermometer readings as the correct classification, the IRHTET misclassified 14 subjects: eight of 95 febrile subjects were misclassified as afebrile (8.4% false negatives) based on a fever cut-off point of 38.0 °C. However, this number was not significantly different from the reference readings ($McNemar \chi^2 = 0.286$, $p = 0.583$). This indicates good agreement between instruments and lack of bias in classification. Of 110 afebrile subjects, six were classified as febrile (5.4% false positives). The sensitivity for the IRHTET was 0.916, and specificity for the IRHTET was 0.945.

A total of 30 participants were misclassified in the TAT group using the reference thermometer classification, and results were significantly different from results with the IRHTET ($McNemar \chi^2 = 16.1$, $p < 0.001$), indicating a lack of agreement between instruments. As shown in Table 3, the IRHTET was less likely than the TAT to misclassify febrile participants as afebrile (based on the reference thermometer readings). Of the 30 misclassified with the TAT, 26 of the 95 febrile participants were
misclassified as afebrile when they were actually febrile (27.4% false negatives). This was significantly different from the IRHTET readings (p < 0.001). The TAT yielded four false positives (3.6%) of 110 afebrile participants, not significantly different from the IRHTET. The sensitivity for the TAT was 0.726, and the specificity was 0.964.

Repeatability analysis

Although this was not a study end point, repeatability of triplicate measurements for the IRHTET and TAT was assessed based on three measurements taken consecutively separated by at least one minute using the formula in the industry standard for thermometers (European Committee...
Ear vs. forehead IR thermometers

![Graph showing mean difference from reference thermometer readings by age group for the febrile subgroup.](image)

**Figure 3** Mean difference from reference thermometer readings by age group for the febrile subgroup.

<table>
<thead>
<tr>
<th>Age group</th>
<th>IRHTET</th>
<th>TAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-24 Mon</td>
<td>-0.58</td>
<td>-0.12</td>
</tr>
<tr>
<td>&gt;2-4.99 Mon</td>
<td>-0.13</td>
<td>-0.07</td>
</tr>
<tr>
<td>&gt;5-10.99 Years</td>
<td>-0.59</td>
<td>-0.67</td>
</tr>
<tr>
<td>&gt;11-19.99 Years</td>
<td>-0.39</td>
<td>-0.28</td>
</tr>
<tr>
<td>&gt;20-29.99 Years</td>
<td>0.67</td>
<td>0.97</td>
</tr>
<tr>
<td>Over 30</td>
<td>0.56</td>
<td>0.56</td>
</tr>
</tbody>
</table>

**Table 3** Classification of febrile status between instruments

<table>
<thead>
<tr>
<th></th>
<th>TAT to reference thermometer</th>
<th>IRHTET to reference thermometer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Afebrile</td>
<td>Febrile</td>
</tr>
<tr>
<td>Afebrile</td>
<td>106</td>
<td>4</td>
</tr>
<tr>
<td>Febrile</td>
<td>26</td>
<td>69</td>
</tr>
<tr>
<td>Total</td>
<td>132</td>
<td>73</td>
</tr>
</tbody>
</table>

The IRHTET was better at detecting fever (at a fever cut-off of ≥38.0 °C) than the TAT when compared to the reference thermometer.

IRHTET, infrared ear thermometer (ThermoScan® PRO 4000 infrared prewarmed tip ear thermometer); TAT, temporal artery thermometer (Temporal Scanner).

Measurements were significantly different from measurements taken with the reference thermometer. Post hoc analyses indicated that the main source of measurement difference was in the febrile population, where mean TAT readings were significantly lower, by −0.42 °C (−0.76 °F), than readings from the reference thermometer. This difference was not seen with the IRHTET, where the mean difference was −0.08 °C (−0.15 °F), which was not significantly different from 0. Mean readings in different age groups did not significantly differ on any measures.

Bland–Altman analyses indicated sufficiently narrow ranges of prediction for all instrument comparisons (Table 2, Fig. 1a,b). In a rank-ordering of interval widths, the confidence intervals were narrower in all comparisons of the IRHTET to the reference thermometer than in comparisons of the TAT with the reference thermometer. Standard deviations for bias in comparing the IRHTET with the reference thermometer were relatively small for this type of comparison.

Both instruments demonstrated repeatability within the acceptance criteria recommended by the European Committee for Standardization (2003). However, the IRHTET demonstrated nearly 1.5 times better repeatability than the TAT.

The IRHTET had a lower rate of misclassification (14 of 205) and demonstrated no differential bias in misclassification (eight false negatives and six false positives). The TAT had a significantly higher rate of misclassification (30 of 205) and demonstrated a differential bias towards false negatives (26 of 30). While the overall odds of misclassifying were more than twofold higher with the TAT than with the IRHTET, the more clinically meaningful fact is the threefold higher rate of false negatives with the TAT. This translated into a sensitivity of 0.726 for the TAT vs. a sensitivity of 0.916 for the IRHTET.

The present study showed that ear thermometers closely track reference measurements during febrile and non-febrile periods in children. A prior study with a prototype of the IRHTET showed that mean difference between the IRHTET and readings from a pulmonary artery catheter was 0.08 °C, with a 95% confidence interval of 0.61 to −0.44 °C (Bock et al. 2005). A comparison of an earlier IRHTET model (without the heated tip) to a reference rectal thermometer (Smirz et al. 2009) showed a small (+0.2 °C) mean bias of the rectal to ear measurement, as would be expected because rectal temperatures are known to be higher than ear measurements. Sensitivity for the IRHTET in this study was 94%, and specificity was 98% in predicting fever based on rectal readings. Nimah et al. (2006) found that ear temperatures are more accurate for Standardization 2003). Because mean temperature is known to differ between the right and left ears (Smirz et al. 2009), all temperatures were taken from the right ear by the same experienced operator. With these precautions in place, repeatability for both instruments was reasonably close to 0.3 °C (0.54 °F), the standard for clinical repeatability. The repeatability was nearly 1.5 times better overall with the IRHTET (0.23 °C) vs. the TAT (0.32 °C).

**Discussion**

The primary end point for this study was a comparison between the relative performance of the IRHTET and the TAT using oral or rectal temperature measurements from a thermistor-based reference thermometer for comparison. Repeated measures GLM analysis indicated that the IRHTET measurements did not differ significantly from the reference thermometer measurements, but the TAT measurements were significantly different from measurements taken with the reference thermometer. Post hoc analyses indicated that the main source of measurement difference was in the febrile population, where mean TAT readings were significantly lower, by −0.42 °C (−0.76 °F), than readings from the reference thermometer. This difference was not seen with the IRHTET, where the mean difference was −0.08 °C (−0.15 °F), which was not significantly different from 0. Mean readings in different age groups did not significantly differ on any measures.

Bland–Altman analyses indicated sufficiently narrow ranges of prediction for all instrument comparisons (Table 2, Fig. 1a,b). In a rank-ordering of interval widths, the confidence intervals were narrower in all comparisons of the IRHTET to the reference thermometer than in comparisons of the TAT with the reference thermometer. Standard deviations for bias in comparing the IRHTET with the reference thermometer were relatively small for this type of comparison.

Both instruments demonstrated repeatability within the acceptance criteria recommended by the European Committee for Standardization (2003). However, the IRHTET demonstrated nearly 1.5 times better repeatability than the TAT.

The IRHTET had a lower rate of misclassification (14 of 205) and demonstrated no differential bias in misclassification (eight false negatives and six false positives). The TAT had a significantly higher rate of misclassification (30 of 205) and demonstrated a differential bias towards false negatives (26 of 30). While the overall odds of misclassifying were more than twofold higher with the TAT than with the IRHTET, the more clinically meaningful fact is the threefold higher rate of false negatives with the TAT. This translated into a sensitivity of 0.726 for the TAT vs. a sensitivity of 0.916 for the IRHTET.

The present study showed that ear thermometers closely track reference measurements during febrile and non-febrile periods in children. A prior study with a prototype of the IRHTET showed that mean difference between the IRHTET and readings from a pulmonary artery catheter was 0.08 °C, with a 95% confidence interval of 0.61 to −0.44 °C (Bock et al. 2005). A comparison of an earlier IRHTET model (without the heated tip) to a reference rectal thermometer (Smirz et al. 2009) showed a small (+0.2 °C) mean bias of the rectal to ear measurement, as would be expected because rectal temperatures are known to be higher than ear measurements. Sensitivity for the IRHTET in this study was 94%, and specificity was 98% in predicting fever based on rectal readings. Nimah et al. (2006) found that ear temperatures are more accurate
measures of core temperature in children than temperatures taken at any other body site (including temple) during both febrile and non-febrile periods. On the other hand, Duru et al. (2012) reported low sensitivity of readings with an ear thermometer in newborns, although specificity was high in this population.

In addition to this study, other studies have also reported greater accuracy with ear thermometers than with temporal artery thermometers (Mangat et al. 2010). Dybwik and Nielsen (2003) found that the temporal thermometer often missed fever; it detected fever in only 33 of 70 participants. Schuh (2004) reported poor correlation between temporal thermometers and rectal core temperature, and Kimberger et al. (2007) also found that a temporal thermometer differed too much from bladder temperature to be useful when highly reproducible core temperature measurements are needed. Langham et al. (2009) likewise found that temporal readings were not a satisfactory substitute for core readings in patients with fever. On the other hand, Lawson et al. (2007) reported that temporal artery temperature measurements agreed closely with the pulmonary artery temperature.

While fever is only one of the parameters used to assess the health of a patient, the presence of fever (or lack thereof) is one of the parameters used to make a diagnosis and determine the need for additional intervention. Clinicians base their decision to administer antipyretics or subject patients to additional blood studies and even order invasive procedures (such as lumbar puncture) based on body temperature. Fever also has prognostic value; several clinical parameters have been shown to correlate with fever. Thermometers are used as decision-making tools to mandate cooling and hydration when the body is subjected to heat stress (athletic events, high ambient temperature conditions). In these situations, an accurate tool for temperature measurement is critical.

For appropriate decision-making, a thermometer should not only be quick and easy to use, but more importantly, it should yield low false-negative and false-positive readings and repeatable, accurate measurements in children and adults in a variety of ambient conditions.

Limitations

In the absence of invasive core body temperature measurements, comparing the readings from the test thermometers to ‘true’ body temperature is not possible. For practical reasons, many studies have used a thermistor-based thermometer reading as a standard in lieu of a true core temperature reading (for example, Braun et al. 1998, Giuliano et al. 2000, Burnham et al. 2006, Kelechi et al. 2006, Mangat et al. 2010) The study results are limited to a comparison with a popular professional-grade, thermistor-based, non-invasive reference thermometer, the SureTemp® Plus electronic contact thermometer.

To shorten the time during which the paediatric participants would have to cooperate and to minimise possible inter-rater variability, the protocol specified that one experienced nurse would take all the temperature measurements. The nurse was right-handed, and taking all measurements from the right in a standardised fashion contributed to the repeatability of the measurements. The disadvantage of having only one rater is the absence of a double check on the measurements.

Conclusions

In conclusion, our data indicate that temporal artery measurements were statistically significantly lower than oral or rectal reference thermometer measurements overall as well as in the febrile category for all ages and that the overall false-negative rate was about threefold higher than with the IRHTET. This suggests that measurements made using the IRHTET more closely reflect measurements from the reference thermometer than those using the TAT in both the febrile and overall population in children of all age groups.

Relevance to clinical practice

The main causes of fever are infections of various types, both viral and bacterial. This is particularly significant in a patient population ranging from 0–24 months. A risk of urinary tract infections, bacteraemia and meningitis are common causes of concern among physicians for young children with high fever. Fever status is frequently used to determine the need for additional intervention (such as lumbar puncture), additional laboratory work (blood/urine tests) and the prescribing of antibiotics and antipyretics. Therefore, it is important to ensure that a thermometer used to assess body temperature will not misclassify fever. Misclassification of fever can be estimated by determining a thermometer’s performance in terms of the false-negative and false-positive rates (also measured in terms of sensitivity and specificity values). The higher the sensitivity values (or lower the false-negative rates), the lower is the chance of misclassifying a fever. This comparison of IR thermometers indicates that the IRHTET ear thermometer is less likely to result in a missed fever than the TAT temporal thermometer.
Acknowledgements

The authors would like to thank Stephanie G. Phillips, PhD, for editorial assistance.

Contributions

Study design: PAH, LSM; data collection and analysis: PAH, MS and manuscript preparation: PAH.

Funding

Funding for this study was provided by Braun, GmbH and Kaz, Inc., Southborough, MA.

Conflict of interest

The authors have no financial interest in the products, technology or methodology discussed in this article.

References


PA Hamilton et al.


The Journal of Clinical Nursing (JCN) is an international, peer reviewed journal that aims to promote a high standard of clinically related scholarship which supports the practice and discipline of nursing.

For further information and full author guidelines, please visit JCN on the Wiley Online Library website: http://wileyonlinelibrary.com/journal/jocn

Reasons to submit your paper to JCN:
High-impact forum: one of the world’s most cited nursing journals, with an impact factor of 1.118 – ranked 30/95 (Nursing (Social Science)) and 34/97 Nursing (Science) in the 2011 Journal Citation Reports® (Thomson Reuters, 2011).
One of the most read nursing journals in the world: over 1.9 million full text accesses in 2011 and accessible in over 8000 libraries worldwide (including over 3500 in developing countries with free or low cost access).
Early View: fully citable online publication ahead of inclusion in an issue.
Fast and easy online submission: online submission at http://mc.manuscriptcentral.com/jcnur.
Positive publishing experience: rapid double-blind peer review with constructive feedback.
Online Open: the option to make your article freely and openly accessible to non-subscribers upon publication in Wiley Online Library, as well as the option to deposit the article in your preferred archive.