Evaluation of Three Brands of Tympanic Thermometer

Carolyn Hoffman, Marion Boyd, Bonnie Briere, Francis Loos, and Peter J. Norton

Infrared tympanic thermometers (ITT) have many documented benefits, including speed, ease of use, and noninvasiveness, to support their use in emergency departments (ED) and intensive care units (ICU). However, concerns have been raised about the accuracy of temperatures reported by ITT. This study was conducted to evaluate the accuracy of 3 brands of ITT, compared to rectal and pulmonary artery thermometers, in ED and ICU.

settings. Results indicated adequate specificity for all 3 ITT in ED (range = 0.9242 to 1.0000) and ICU (range = 0.9737 to 1.0000), but unacceptable sensitivity in ED (range = 0.5455 to 0.8000) and ICU (range = 0.0000). Further analyses indicated highly variable ITT-reported temperatures. In ED, ITT temperatures were, on average, 0.3684°C lower, and could be expected to vary by more than 2°C from the actual temperature as reported by rectal thermometer. In ICU, the average reported temperature was similar to the actual pulmonary artery temperature (0.0259°C lower), but again could vary by more than 2°C. Recommendations for nursing and education are discussed.

Literature Review

Speed, ease of use, and noninvasiveness are documented benefits of the infrared tympanic thermometer (ITT) (Alexander & Kelly, 1991a, 1991b; Erickson & Meyer, 1994; Klein et al., 1993; Shinozaki, Deane, & Perkins, 1988; Yaron, Lowenstein, & Koziol-McLain, 1995). However, concerns have been raised about its accuracy (Chamberlain et al., 1995; Davis, 1993; Edge & Morgan, 1993; Erickson & Kirklin, 1993; Jakobsson, Nilsson, & Carlsson, 1992; Milewski, Ferguson, & Terndrup, 1991; Talo, Macknin, & VanderBrug-Medendorp, 1991; Zehner & Terndrup, 1991).

In previous studies, the ability to detect clinically significant fevers using the ITT was found to be a controversial issue (Brennan, Falk, Rothrock, & Kerr, 1995; Rhodes & Grandner, 1990). Factors contributing to conflicting results include size of auditory canal, especially with patients under 36 months of age (Klein et al., 1993), operator technique (e.g., ear tug, cleaning of probe tip) (Erickson & Meyer, 1994; Lattavo, Brit, & Dobal, 1995; Nobel, 1992; Terndrup & Rajk, 1992), environmental and ambient temperatures (Chamberlain et al., 1991; Doyle, Zehner, & Terndrup, 1992; Thomas, Savage, & Bregelmann, 1997; Zehner & Terndrup, 1991), model of thermometer (Erickson & Meyer; Klein et al.; Lattavo et al.), operating modes and mathematical corrective values used (Brennan et al.; Fraden & Lackey, 1991; Romano et al., 1993; Schmitz, Blair, Falk, & Levine, 1995), occlusion of ear canal with cerumen (Romano et al.; Yaron et al., 1995), and otitis media in children (Kelly & Alexander, 1991; Romano et al.).

We noted that there were incomplete data on: (a) the newer generations of ITT, (b) the need for repair and recalibration, and (c) ease of use (Nobel, 1992), and we therefore decided to address some of these issues within our clinical practice settings. The purpose of this study was to compare the three ITT models to determine their level of agreement with rectal or pulmonary artery (PA) temperature measurements.
Tympanic Thermometers

Methods

A prospective study was undertaken with clients in the emergency department (ED) of the Pasqua Hospital and the intensive care units (ICU) of the Pasqua Hospital and the Plains Health Centre. Both institutions are within the Regina Health District in the province of Saskatchewan. The study took place over a period of 11 months.

Subjects

The subjects chosen had routine rectal (ED) or pulmonary artery (ICU) temperature assessment and were at least 3 months of age (Chamberlain et al., 1991; Davis, 1993; Stewart & Webster, 1992). Excluded were clients who had pre-existing rectal anomalies or surgeries that prevented rectal temperature assessment (Yaron et al., 1995), bilateral ear pain (Yaron et al.), or suspected cerebrospinal fluid draining from the ear, as were pediatric clients under the care of the Allan Blair Cancer Centre.

A total of 304 clients (46.7% female) treated in the ED had temperature recorded using rectal thermometer and ITT. A total of 108 clients (32.4% female) treated in ICU had temperature recorded using PA thermometer and ITT. ED clients ranged in age from 3 months to 87 years (mean = 3.8 years, SD = 11.995). ICU clients ranged in age from 40 years to 84 years (mean = 66.8 years, SD = 9.860).

Thermometers

Five types of thermometer were used. The RT and PAT thermometers were selected to be used as benchmarks in comparing the agreement and accuracy of the three ITT. The assertion of Romano et al. (1993) that “pulmonary artery blood temperature is traditionally accepted as the reference measurement of core body temperature” (p. 1181) is supported by other investigators (Ferrara-Love, 1997; Jakobsson et al., 1992; Neirman, 1991). Table 1 lists the brand, manufacturer, and number of measurements for ITT, rectal thermometer, and PA thermometer.

Each ITT model was submitted to Clinical Engineering for initial calibration. All models were found to be within 0.1°C of their respective set points, well within the specifications for these models. No repair or recalibration was required for any model during the data-collection periods. Twelve nurses were trained in the recommended procedure for obtaining ITT measurement, either by company representatives or by means of a video.
<table>
<thead>
<tr>
<th>Method</th>
<th>ED n</th>
<th>ICU n</th>
<th>Instrument</th>
<th>Manufacturer</th>
<th>Collection Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITT</td>
<td>101</td>
<td>28</td>
<td>IVAC Core•Check (Model 2090)</td>
<td>IVAC Corp., San Diego, California</td>
<td>23/10/95 to 9/11/95</td>
</tr>
<tr>
<td>ITT</td>
<td>102</td>
<td>39</td>
<td>First Temp Genius (Model 3000A)</td>
<td>Intelligent Medical Systems, Carlsbad, California</td>
<td>30/11/95 to 13/01/96</td>
</tr>
<tr>
<td>ITT</td>
<td>101</td>
<td>41</td>
<td>Thermoscan Pro-1 Instant Thermometer (Model IR-1)</td>
<td>Thermoscan Inc., San Diego, California</td>
<td>30/04/96 to 29/06/96</td>
</tr>
<tr>
<td>Rectal</td>
<td>304</td>
<td>0</td>
<td>IVAC Temp*Plus II (Model 2080A)</td>
<td>IVAC Corp., San Diego, California</td>
<td>Above periods</td>
</tr>
<tr>
<td>PAT</td>
<td>0</td>
<td>108</td>
<td>Baxter Swan-Ganz® 7F Thermodilution Catheter (connected to Hewlett Packard Cardiac Output Model M1012A)</td>
<td>Baxter Health Care Corporation, Irvine, California</td>
<td>Above periods</td>
</tr>
</tbody>
</table>
Tympanic Thermometers

Procedure

Oral consent was obtained from the client or family. Written consent was not required, as ITT is noninvasive and atraumatic and the sample had RT or PAT taken routinely. After a waiting time of 5 minutes in the ED, the RT was taken by lubricating the probe and placing it in the rectum until the completion tone was heard. PAT was sensed at the tip of the catheter, which sits in the pulmonary artery, and transmitted to a monitor. Each of the nurses obtained a tympanic temperature within 1 minute of obtaining RT or PAT. Where possible, nurses used their dominant hand to obtain the tympanic measurement from the ipsilateral ear of the client (e.g., right hand dominant: right ear of client), as recommended by Yaron et al. (1995).

Data collected for each client included age, sex, date, time of temperature assessment, whether right or left ear, temperature values obtained, diagnosis of client, any other relevant conditions (e.g., time of arrival in emergency), and the initials of the nurse data collector. In addition, a tympanic thermometer evaluation form was provided to all data collectors after each ITT model had been used. Completion of the form was voluntary and the information remained confidential.

Results

The data were analyzed using SPSS (Statistical Package for Social Sciences) for Windows. Descriptive temperature data are provided in Table 2.

Sensitivity and Specificity

The data from the three tympanic models were analyzed to determine their sensitivity (Sens.) and specificity (Spec.) in distinguishing fever and no-fever, as classified by rectal or PA thermometer. As stated by Tuokko and Hadjistavropoulos (1998), “the sensitivity of a test at any given cut-off score is the proportion of [fever positive] persons with scores above the test positive range. The specificity of a test score is the proportion of [fever negative] persons with scores falling in the test negative range” (pp. 25–26). Sensitivity and specificity analyses were not conducted for hypothermia, as too few cases of hypothermia presented in either ED (n = 2) or ICU (n = 3) to allow us to draw solid conclusions.

The ITT held very strong specificity, correctly identifying as fever-negative (temperature ≤38.5°C) clients who were fever-negative as determined by rectal thermometer. Specifically, IVAC perfectly identified all fever-negative clients as fever-negative (Spec. = 1.0000), while
<table>
<thead>
<tr>
<th>Unit/Method</th>
<th>Brand</th>
<th>Aggregate</th>
<th>Thermoscan ($n=101$)</th>
<th>Genius ($n=102$)</th>
<th>Thermoscan ($n=41$)</th>
<th>Genius ($n=39$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(N=304)</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SD</td>
<td>SD</td>
<td>SD</td>
<td>SD</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>37.84</td>
<td>38.07</td>
<td>37.09</td>
<td>37.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.13</td>
<td>1.2</td>
<td>0.93</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.13</td>
<td>0.18</td>
<td>ICU/PAT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.6</td>
<td>0.33</td>
<td>ICU/PAT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.6</td>
<td>ICU/PAT</td>
<td>Mean</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ICU/PAT</td>
<td></td>
<td>Difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.04</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: ITT Mean Temperatures Compared to RI and PAT
Genius (Spec. = 0.9870) and Thermoscan (Spec. = 0.9242) were nearly as accurate. The specificity of the ITT in determining fever was equally accurate with the PA thermometer. IVAC and Thermoscan perfectly identified all fever-negative clients as fever-negative (Sens. = 1.0000), while Genius was nearly as accurate (Spec. = 0.9737). In sensitivity, however, the ITT was less than ideal. Of the ED clients who were fever-positive (>38.5°C) as determined by RT thermometer, both Genius and Thermoscan identified 80% (Sens. = 0.8000), while the IVAC correctly identified only about 50% (Sens. = 0.5455). In the ICU, none of the ITT correctly identified a fever-positive client (Sens. = 0.0000), although by coincidence no clients presenting with fever (as determined by PAT) were tested with Thermoscan.

Agreement

The degree of agreement between temperatures reported by ITT and reported by rectal and PA thermometers was tested using the Bland and Altman technique (Bland & Altman, 1986; Szafarski & Slaughter, 1996; Yaron et al., 1995). This technique overcomes the methodological problems associated with using correlation analyses for determining agreement between clinical measures of the same parameter (Nield & Gocka, 1993; Szafarski & Slaughter).

The ED data, using RT as the comparison measure, indicated that on average the ITT-reported temperatures were 0.3684°C lower than those reported by rectal thermometer (see Figure 1). Furthermore, calculating 95% confidence intervals (95% CI) showed that ITT-reported temperatures could be expected to range from over-reporting (i.e., ITT higher than RT) by 0.6796°C (lower level of agreement) to under-reporting (i.e., ITT lower than RT) by 1.4164°C (upper level of agreement).

The ICU data, using PAT as the comparison measure, showed a more accurate average (see Figure 2), under-reporting by 0.0259°C. Calculation of 95% CI, however, revealed an expected range from 1.0623°C over-reported (lower level of agreement) to 1.1141°C under-reported (upper level of agreement).

Differences

Differences in individual temperatures were also examined for patterns. Difference tables were generated to determine the level of agreement among assessment instruments.

The IVAC temperature was lower than the RT in 85% of clients, and the difference was more than 0.5°C in 58% of clients. The Genius tem-
Figure 1  Scatterplot of Difference between ITT and Rectal Temperatures (in ED)

Figure 2  Scatterplot of Difference between ITT and PA Temperatures (in ICU)
temperature was lower than the RT in 64% of clients, and the difference was more than 0.5°C in 34% of clients. The Thermoscan temperature was lower than the RT in 66% of clients, but the under-reporting was more than 0.5°C in only 26% of clients.

The IVAC temperature was lower than the PAT in 53% of clients, and the difference was more than 0.5°C in 11% of clients. Genius, while having the most temperatures exactly equal to the PAT (10%), also reported a lower temperature than the PAT in 47% of clients, by more than 0.5°C in 26% of clients. On the other hand, Thermoscan, also with temperatures exactly equal to the PAT in 10% of clients, recorded temperatures higher than the PAT in 63% of clients, although higher by more than 0.5°C in only 12% of clients. Thermoscan recorded temperatures lower than the PAT in 27% of clients.

In Table 3, the difference data are further broken down by age for the rectal group. On the whole, Thermoscan reported the closest temperature to the RT in all age groups.

**User Evaluation**

The comments related to each brand of ITT with respect to ease of use, accuracy, speed of temperature determination, probe covers, cleaning, sturdiness, battery performance, maintenance, and recommendations revealed that overall the Genius and Thermoscan models had very positive evaluations. They were found to be lightweight, quick, easy to use, easy to clean, and easy to maintain. Overall approval for both units was very high, and some data collectors recommended purchasing the units. On the other hand, although IVAC received some positive comments on ease of use, speed of temperature determination, and sturdiness, for most individuals completing the evaluation form accuracy was a serious concern. Overall approval for IVAC was very low and purchase was not recommended.

**Discussion**

The purpose of this study was to test tympanic thermometry against both the gold standard of the rectal thermometer (ED group) and the PA thermometer (ICU group). We sought to determine whether the level of agreement would be high enough to allow us to simply accept tympanic temperature assessment as an alternative to other measures. Finally, we hoped to identify one model of tympanic thermometer that showed better agreement than the others and that therefore could be recommended for purchase.
<table>
<thead>
<tr>
<th>ITT Brand</th>
<th>Age</th>
<th>&lt; -0.5°C</th>
<th>-0.5°C to 0°C</th>
<th>0°C</th>
<th>0 to +0.5°C</th>
<th>&gt; +0.5°C</th>
<th>% (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IVAC</td>
<td>3–12 mos</td>
<td>5 (1)</td>
<td>10 (2)</td>
<td>5 (1)</td>
<td>35 (7)</td>
<td>45 (9)</td>
<td>100 (20)</td>
</tr>
<tr>
<td></td>
<td>12–36 mos</td>
<td>1 (1)</td>
<td>7 (5)</td>
<td>3 (2)</td>
<td>27 (19)</td>
<td>62 (43)</td>
<td>100 (70)</td>
</tr>
<tr>
<td></td>
<td>&gt;36 mos</td>
<td>9 (1)</td>
<td>9 (1)</td>
<td>9 (1)</td>
<td>64 (7)</td>
<td>100 (11)</td>
<td></td>
</tr>
<tr>
<td>Genius</td>
<td>3–12 mos</td>
<td>0</td>
<td>18 (7)</td>
<td>0</td>
<td>33 (13)</td>
<td>49 (19)</td>
<td>100 (39)</td>
</tr>
<tr>
<td></td>
<td>12–36 mos</td>
<td>5 (3)</td>
<td>22 (13)</td>
<td>9 (5)</td>
<td>34 (20)</td>
<td>30 (17)</td>
<td>100 (58)</td>
</tr>
<tr>
<td></td>
<td>&gt;36 mos</td>
<td>20 (1)</td>
<td>40 (2)</td>
<td>0 (0)</td>
<td>20 (1)</td>
<td>20 (1)</td>
<td>100 (5)</td>
</tr>
<tr>
<td>Thermoscan</td>
<td>3–12 mos</td>
<td>3 (1)</td>
<td>19 (6)</td>
<td>3 (1)</td>
<td>50 (18)</td>
<td>25 (9)</td>
<td>100 (35)</td>
</tr>
<tr>
<td></td>
<td>12–36 mos</td>
<td>6 (4)</td>
<td>28 (16)</td>
<td>9 (5)</td>
<td>43 (25)</td>
<td>14 (8)</td>
<td>100 (58)</td>
</tr>
<tr>
<td></td>
<td>&gt;36 mos</td>
<td>0</td>
<td>25 (2)</td>
<td>12.5 (1)</td>
<td>50 (4)</td>
<td>12.5 (1)</td>
<td>100 (8)</td>
</tr>
</tbody>
</table>
Tympamic Thermometers

Initial analyses indicated strong specificity for all ITT models compared to either rectal or PA thermometers. Thus, there was a low incidence of ITT reporting fever in patients without fever. Conversely, however, ITT failed to detect fever in a clinically significant number of patients with fever. As treatment decisions are often based on the presence or absence of fever, the poor ITT sensitivity could potentially result in serious clinical implications.

Assessment of the agreement between ITT and established thermometers (rectal and PA) was not supportive of the continued use of ITT. Although the mean difference between ITT and either rectal or PA temperature did not differ to a vast degree, the variability is cause for concern. This study revealed that ITT-reported temperatures could be expected to vary upwards of 1°C from the patient’s actual temperature. In our clinical setting, this degree of error is unacceptable.

A possible limitation of the study, and a possible explanation for the discrepancies found between ITT and established instruments, is measurement error, which is possible with any type of thermometer. The site of measurement is known to be a factor, because of the different tissue — the goal being to measure core temperature. In addition, how the thermometer is used, its speed of calibration to the surrounding temperature, the ambient temperature, and the presence or absence of various body excreta or secretions may also cause measurement error. We made no attempt to account for these factors. Finally, as with all methods, the potential for operator error may confound the measurement. A further limitation of our study was the use of different training methods for each model of thermometer, based on the manufacturer’s recommendation. In addition, our study may have been influenced by the use of a single model of thermometer at a time. No attempt was made to randomize the order of the type of ITT used for data collection. Any or all of these factors may account for the difference between temperatures taken at two different sites using different types of technology.

The potential for measurement error suggests the need for repeated measures using all implements if there is any suspicion that a temperature reading is incorrect or if a temperature reading approaches the set diagnostic points for hypothermia or hyperthermia. However, the low sensitivity of the ITT compared to standard measures makes its clinical use questionable. The data from this study suggest that the ITT at its best failed to detect fever in one out of five clients. The client risks associated with this degree of error are self-evident.
Recommendations for Nursing

In general, clients who require temperature assessment should have their temperature taken by standard rectal or oral thermometry. In the select population in which a PA catheter is in place, PAT is the most accurate measurement (Ferrara-Love, 1997; Neirman, 1991). Rectal and oral methods have the advantage of cost efficiency, as these probe covers are less expensive than tympanic covers. In specific populations in which there is no PA and/or in which rectal or oral assessment is inappropriate, such as the unconscious or anesthetized client or the young child with contraindications to RT measurement, ITT is worth considering (Erickson & Yount, 1991).

Education in the use of tympanic thermometry should include the potential for ITT disagreement with other measures. In situations in which comparative core temperature is required or in which inaccuracy is suspected, staff should be educated in the need for repeated measures to identify false measurements. In addition, as temperatures reach diagnostic set points, staff should be encouraged to repeat measures or add a second measure of temperature.

Conclusions

Infrared tympanic thermometers may have many advantages for both clients and health-care professionals. This study, however, raises questions regarding the agreement of ITT with other assessment methods. It is our opinion, based on the results of this study, that ITT requires further research and development before it is used as the procedure of choice. ITT can be used with caution in situations where other methods of temperature assessment are contraindicated.

In closing, the data from this study allow us to recommend the Thermoscan, because of its relatively close agreement with actual RT or PAT readings, for use in specific client populations. Certainly, in the absence of a PA, or in the case of an unconscious or uncooperative client who is at risk of rectal perforation, oral damage, or a broken probe, the tympanic thermometer remains an option.

References

Tympanic Thermometers


Correspondence concerning this article should be addressed to Francis Loos, Surgical Intensive Care Unit - 2A, Regina General Hospital, 1440-14th Avenue, Regina, SK S4P 0W5. Telephone: 306-766-3974. Fax: 306-766-3978. E-mail: floos@reginahealth.sk.ca