



# Centre for Clinical Effectiveness

## Evidence Request 2007-005b

### Factors affecting the accuracy of infrared tympanic thermometers and digital/electronic or infrared axillary temperature recordings in children

Date: 6 June 2007

Authors: Jane Reid, Tari Turner and Claire Harris

Requestors: Nicole Lloyd, Nurse Educator, Emergency Department, Casey Hospital

#### Abstract

**Background:** The nurse educator in the emergency department of Casey Hospital requested evidence regarding whether infrared tympanic thermometers or axillary thermometers are more accurate at measuring children's temperatures.

She also requested information about the factors that affect the accuracy of these two types of thermometers. This report addresses this question. Please refer to Evidence Request 2007-005a for the first question.

**Clinical Questions:**

1. In children and infants having temperature assessed by infrared tympanic thermometer, what factors affect the accuracy of temperature recording?
2. In children and infants having temperature assessed by digital/electronic or infrared axillary thermometer, what factors affect the accuracy of temperature recording?

**Methods:** We included all trials published in English and conducted on humans.

We searched All EBM Reviews (Ovid) (including The Cochrane Database of Systematic Reviews, The Cochrane Central Register of Controlled Trials, the Database of Abstracts of Reviews of Effects (DARE), ACP Journal Club) and the Cochrane library for Health Technology Assessments. We also searched Medline and CINAHL and several key guideline websites.

Studies were selected and appraised by one reviewer in consultation with colleagues using inclusion, exclusion and appraisal criteria established *a priori*.

**Results:** 1. Infrared tympanic thermometry

There were a total of 588 references found and a systematic review and seven cohort studies were reviewed. These eight articles evaluated the factors: mode of device, model of device, child's age, presence of otitis media, use of ear tug, presence of cerumen, and nurse taking the measurements. Only mode of device was found to have any effect on accuracy of infrared tympanic thermometer measurements; the effect found was clinically important but not statistically significant. None of the other factors were found to have any effect on accuracy.

2. Axillary thermometry

There were a total of 130 references found and two cohort studies were reviewed. These articles evaluated mode of device and child's age. Neither of these factors were found to effect the accuracy of digital/electronic or infrared axillary thermometry.

**Conclusions:** 1. Infrared tympanic thermometry

Although there were no statistically significant differences found between the modes used, there were some clinically important differences. The smallest differences were seen when using the infrared tympanic thermometer in rectal and core modes. That is to say these modes provided the most accurate result. When actual and tympanic modes were used there were the largest differences; these modes provided the least accurate result. Model of device, child's age, presence of otitis media and ear tug do not appear to affect the accuracy of the tympanic temperature measurements. There is insufficient evidence to draw any conclusions about how the presence of cerumen or the nurse taking the measurements affects the accuracy of thermometer measurement.

2. Axillary thermometry

There is insufficient evidence to draw any conclusions about how the mode of device or child's age affects the accuracy of temperature measurements.

## Background

The nurse educator in the emergency department of Casey Hospital requested evidence regarding whether infrared tympanic thermometers or axillary thermometers are more accurate at measuring children's temperatures.

She also requested information about the factors that effect the accuracy of these two types of thermometers. This report addresses this question. Please refer to Evidence Request 2007-005a for the first question.

## Clinical Question

1. In children and infants having temperature assessed by infrared tympanic thermometer, what factors affect the accuracy of temperature recording?
2. In children and infants having temperature assessed by digital/electronic or infrared axillary thermometer, what factors affect the accuracy of temperature recording?

## Methods

### Study Selection Criteria

#### Inclusion and exclusion criteria

Patient	<p>1. Infrared tympanic thermometer</p> <p>Inclusion:</p> <ul style="list-style-type: none"> <li>• Children and infants (0-18 years), having temperature assessed by infrared tympanic thermometer.</li> </ul> <p>Exclusion:</p> <ul style="list-style-type: none"> <li>• Preterm babies (&lt;37 weeks gestational age)</li> <li>• Infants in heated bassinets, cribs, incubators or radiant warmers</li> <li>• A population of solely hypothermic or hyperthermic participants</li> <li>• Children being cooled, warmed or rewarmed</li> <li>• Cranio-facial or ear canal malformations</li> <li>• All other (non-infrared) tympanic thermometers</li> </ul>	<p>2. Digital/electronic or infrared axillary thermometer</p> <p>Inclusion</p> <ul style="list-style-type: none"> <li>• Children and infants (0-18 years), having temperature assessed by digital/electronic or infrared axillary</li> </ul> <p>Exclusion:</p> <ul style="list-style-type: none"> <li>• Preterm babies (&lt;37 weeks gestational age)</li> <li>• Infants in heated bassinets, cribs, incubators or radiant warmers</li> <li>• A population of solely hypothermic or hyperthermic participants</li> <li>• Children being cooled, warmed or rewarmed</li> <li>• Cranio-facial or ear canal malformations</li> <li>• All other axillary thermometers (ie glass-mercury)</li> <li>• Axillary thermometers which use a skin temperature sensor</li> </ul>
Comparison factors	<p>Inclusion:</p> <ul style="list-style-type: none"> <li>• Any factors that may impact on the accuracy of temperature recording, including (but not limited to) child's age, assessor's clinical expertise, otitis media, ear tug etc.</li> </ul> <p>Exclusion:</p> <ul style="list-style-type: none"> <li>• Nil</li> </ul>	
Outcomes	<p>Inclusion:</p> <ul style="list-style-type: none"> <li>• Temperature</li> </ul> <p>Exclusion:</p> <ul style="list-style-type: none"> <li>• All other outcomes</li> </ul>	
Reference standard	Core body temperature (arterial blood, bladder or rectal) using digital/electronic or glass-mercury thermometer	
Study design	<p>Inclusion:</p> <ul style="list-style-type: none"> <li>• Prospective cohort studies where infrared tympanic and core body temperature have been measured in all participants.</li> <li>• Systematic review of the above study type</li> </ul> <p>Exclusion:</p> <ul style="list-style-type: none"> <li>• All other studies</li> </ul>	<p>Inclusion:</p> <p>Prospective cohort studies where digital/electronic or infrared axillary and core body temperature have been measured in all participants.</p> <p>Systematic review of the above study type</p> <p>Exclusion:</p> <ul style="list-style-type: none"> <li>• All other studies</li> </ul>
Publication details	<p>Inclusion: Studies in English and conducted on humans</p> <p>Exclusion: Languages other than English</p>	

## Search Strategy

Evidence Source	Date of Search
<u>Databases:</u> <ul style="list-style-type: none"> <li>All EBM Reviews (Ovid) *</li> <li>Medline (Ovid) (from 1966)</li> <li>CINAHL (Ovid) (from 1982)</li> <li>The Cochrane Library**</li> </ul>	13 <sup>th</sup> February 2007 13 <sup>th</sup> February 2007 13 <sup>th</sup> February 2007 13 <sup>th</sup> February 2007
<u>Relevant websites</u> <ul style="list-style-type: none"> <li>The Joanna Briggs Institute: <a href="http://www.joannabriggs.edu.au">www.joannabriggs.edu.au</a></li> <li>US National Guideline Clearinghouse: <a href="http://www.guidelines.gov">www.guidelines.gov</a></li> <li>UK National Institute for Clinical Excellence: <a href="http://www.nice.org.uk">www.nice.org.uk</a></li> <li>Scottish Intercollegiate Guidelines: <a href="http://www.sign.ac.uk">www.sign.ac.uk</a></li> <li>New Zealand Guidelines Group: <a href="http://www.nzgg.org.au">www.nzgg.org.au</a></li> <li>NH&amp;MRC guidelines <a href="http://www.nhmrc.gov.au">www.nhmrc.gov.au</a></li> <li>Guidelines International Network: <a href="http://www.g-i-n.net">www.g-i-n.net</a></li> <li>Trip Database: <a href="http://www.tripdatabase.com">www.tripdatabase.com</a></li> <li>Google: <a href="http://www.google.com.au">www.google.com.au</a></li> </ul>	21 <sup>st</sup> Nov 2006 21 <sup>st</sup> Nov 2006 21 <sup>st</sup> Nov 2006 21 <sup>st</sup> Nov 2006 21 <sup>st</sup> Nov 2006 21 <sup>st</sup> Nov 2006 21 <sup>st</sup> Nov 2006 21 <sup>st</sup> Nov 2006 21 <sup>st</sup> Nov 2006 21 <sup>st</sup> Nov 2006

\*(including The Cochrane Database of Systematic Reviews, The Cochrane Central Register of Controlled Trials, The Database of Abstracts of Reviews of Effects (DARE), ACP Journal Club)

\*\* for Health Technology Assessments

### Search Terms Q1

<b>Patient</b>	exp Child/ OR exp Pediatrics/ OR exp Infant/ OR (child\$ or pediatric\$ or paediatric\$ or infant\$).mp AND ((thermomet\$ or temperature\$).mp. OR exp Body temperature/) AND (tympanic\$ or ear\$ or aural).mp.
<b>Comparison Factors</b>	-
<b>Outcomes</b>	(accura\$ or precis\$ or sensitiv\$ or specific\$ or valid\$ or reliabl\$ or predict\$).mp. OR exp "Sensitivity and Specificity"/ OR exp "Reproducibility of Results"/

### Search Terms Q2

<b>Patient</b>	exp Child/ OR exp Pediatrics/ OR exp Infant/ OR (child\$ or pediatric\$ or paediatric\$ or infant\$).mp AND ((thermomet\$ or temperature\$).mp OR exp Body temperature/) AND Axilla\$.mp
<b>Comparison Factors</b>	-
<b>Outcomes</b>	(accura\$ or precis\$ or sensitiv\$ or specific\$ or valid\$ or reliabl\$ or predict\$).mp OR exp "Sensitivity and Specificity"/ OR exp "Reproducibility of Results"/

## Data Collection & Analysis

Studies were selected and appraised by one reviewer in consultation with colleagues, using inclusion, exclusion and appraisal criteria established *a priori*.

In the articles reviewed, temperature measurements were usually reported in degrees Centigrade. In articles in which temperature was reported in Fahrenheit, these have been converted into Centigrade and both the readings documented (using <http://www.online-calculators.co.uk/conversion/fahrenheitcentigrade.php>). The Centigrade figures appear in italics and in brackets.

### Definitions of terms used

- Accuracy.** In this review, accuracy refers to the closeness of the test measurement to the reference standard; that is the difference between the reference standard and the test measurement. The lower the figure, the more accurate the device is.
- Area under the curve (AUC).** See Receiver Operating Characteristic (ROC) curve analysis

- **Ear tug.** A technique used, sometimes recommended by the manufacturer, to straighten the external ear canal and allow unobscured access to the tympanic membrane<sup>4</sup>.
- **Mode** Depending on the mode selected, a numerical constant, known as an offset, is added to the measured temperature to generate estimated core, tympanic, oral, or rectal temperature with which clinicians may be more familiar<sup>1</sup>.

In addition to the above modes, which relate to sites in the body, electronic thermometers offer two options for measuring body temperature: 1) the predictive mode uses a predictive algorithm to calculate temperature in approximately 60 seconds, 2) The monitor mode continuously measures the patient's temperature, taking up to three minutes to achieve a stable temperature<sup>9</sup>.

- **Receiver Operating Characteristic (ROC) plot.** A receiver operating characteristic plot is obtained by calculating the sensitivity and specificity of every observed data value and plotting sensitivity against 1 – specificity<sup>15</sup>. A global assessment of the performance of the test (sometimes called diagnostic accuracy) is given by the area under the receiver operating characteristic curve<sup>15</sup> (AUC). The closer this is to 1.0 the more accurate the test is.
- **Sensitivity.** Sensitivity is the proportion of true positives that are correctly identified by the test<sup>14</sup>. For example, if there were 100 cases of fever (by rectal thermometer) and the tympanic thermometer detected 70 of these; that would be a sensitivity of 70%. The higher number the better able the test is to detect fever.
- **Specificity.** Specificity is the proportion of true negatives that are correctly identified by the test<sup>14</sup>. For example, if there were 100 cases of non-fever (by rectal thermometer) and the tympanic thermometer detected 30 of these; that would be a sensitivity of 30%. The higher number the better able the test is to detect non-fever.

## Results

### 1. Factors which affect the accuracy of infrared tympanic thermometry

There were a total of 588 references found with the above search terms. After reviewing the title, abstract or full text, one systematic review<sup>1</sup> and eleven cohort studies<sup>2-8,10-13</sup> were found which addressed the research question. Four cohort studies were excluded because they undertook insufficient or inappropriate data analysis<sup>10-13</sup>.

The one systematic review and seven cohort studies assessed the factors: mode of device, model of device, child's age, presence of otitis media, use of ear tug, presence of cerumen and nurse taking measurements. Results of the included studies are summarised here. Full details of appraisals can be found in Appendix 1.

#### Mode of device

One systematic review published in 2002 was found which examined the impact of mode of the device.

The review by Craig et al<sup>1</sup> included 25 studies and 4441 children (although data reported on only 4344). The review included method comparison studies in which temperature taken using an infrared tympanic thermometer (ITT) was compared to rectal thermometry (electronic, glass-mercury or indwelling thermocouple device) and which had included children between 0 and 18 years. Studies were excluded that enrolled preterm neonates (born at <37 weeks gestation) or hypothermic only participants (rectal temp <35.0°C).

This review was of high quality. Although, the validity of trials was appraised, the assessors used limited criteria and did not discuss the strengths and limitations of the studies in great detail.

The review found that there were no statistically significant differences between the modes used. We, the authors of this review, note that there were some clinically important differences. The smallest differences (rectal temperature measurement minus ITT measurement) were seen when the ITT was used in rectal or core modes. When actual or tympanic modes were used there were the greatest differences. Pooled mean difference (rectal minus tympanic) were: rectal mode 0.15°C (SD 0.55°C); actual mode 0.70°C (SD 0.45°C); core mode 0.25°C (SD 0.51°C); oral mode 0.34°C (SD 0.60°C); tympanic mode 0.62°C (SD 0.51°C); mode not stated 0.32°C (SD 0.45°C).

#### Model of device

Three studies were found which addressed this factor.

Erickson and Woo<sup>3</sup> conducted a study in 1992 of 30 children in intensive care and cardiac recovery units who required a bladder catheter for their care. These children were aged 1 to 45 months of age, (mean of 16.6 months ± standard deviation (SD) 12.6 months).

Three infrared thermometers models were used: A) the Core-Check model 2090A, B) First Temp Genius model 3000A, and C) Thermoscan Pro-1 model IR-1. The devices were set on core mode. Two readings were made with each thermometer, one using ear tug and the other without ear tug. All six readings were made in the same ear in random order, at least one minute apart to allow recovery from transient tissue cooling due to contact with the comparatively cool probe. The infrared tympanic thermometer units were calibrated before, during and after data collection. The reference standard was bladder temperature measurement and these were continuously monitored using: 1) Mon-a-therm Temperature Monitoring System model 4070 (thermistor), or 2) Mon-a-therm CathTemp used

with Mon-a-therm Temperature Monitoring System model 6510 (thermocouple). A rectal (thermocouple) was also used.

This study was of moderate quality. The study subjects were a convenience sample who required bladder catheter; no other inclusion or exclusion criteria have been documented. The participants had six ear measurements taken, and it appears for each of these there was a corresponding axillary, bladder and rectal temperature taken (six sets of measurements (24 measurements) on each child). Taking multiple sets of measurements on individual children may falsely reduce the variability, which may lead to a reduction in the generalisability of the results. All temperature measurements were taken by the same person, therefore these could not be independent or blind. It does appear these were objective as the readings were taken from monitor displays.

The mean difference (test measurement minus bladder measurement) of the different models was: model A  $-0.68^{\circ}\text{C}$  (SD  $0.41^{\circ}\text{C}$ ), model B  $-0.29^{\circ}\text{C}$  (SD  $0.41^{\circ}\text{C}$ ) and model C  $-0.46^{\circ}\text{C}$  (SD  $0.50^{\circ}\text{C}$ ). The bladder measurements were higher. Analysis of variance showed that these differences were statistically significant ( $p < 0.01$ ) and post-hoc testing indicated that model A's values were significantly more negative (less close to bladder temperature) than those with model B ( $p < 0.01$ ) and also significantly more negative than those of model C's ( $p = 0.04$ ). That is to say model B was the most accurate, then model C. Model A was shown to be the least accurate.

Sensitivity of the tympanic thermometers to detect fever ranged from 65% to 92%; specificity from 81% to 91%. (Sensitivity for model A 65%, B 75% and C 92%. Specificity for model A 91%, B 81%, C 86%). Fever was defined as bladder temperature  $\geq 38^{\circ}\text{C}$ , model A  $\geq 37.3^{\circ}\text{C}$ , model B  $\geq 37.7^{\circ}\text{C}$ , model C  $\geq 37.5^{\circ}\text{C}$ . The study used different cut-off points for fever because the average differences (test measurement minus bladder measurement) of each device. Model C does appear clinically better able to differentiate between fever and non-fever than model A or model B, but as no statistical testing or confidence intervals were reported, the statistical significance is unclear.

Taking all the above information into account, it does appear that model B (First Temp Genius model 3000A) and model C (Thermoscan PRO-1 model IR-1) were similar in accuracy and model A (Core-check model 2090A) was least accurate.

Romano et al<sup>6</sup> conducted a study between 1991 and 1992 of 20 children in an intensive care unit of a tertiary paediatric hospital requiring pulmonary artery (PA) catheter monitoring. The median age was three years and seven months (range six months to 15 years). Exclusion criteria were the inability to visualise tympanic membrane, presence of tympanostomy tubes or contraindication to the use of indwelling rectal tubes.

The study tested two types of ITT, the Genius First Temp and Thermoscan Pro-1. Each unit was maintained in core mode. An ear tug manoeuvre was used for the latter machine only, as recommended by the manufacturer. The reference measurement was taken using an Opticath PA catheter inserted using standard procedures. The study also used a Mon-A-Therm rectal probe. Measurements were taken 'simultaneously' by one of the authors.

This study was of moderate quality. All temperature measurements were taken by the same person, therefore these could not be independent or blind. It does appear these were objective as the readings were taken from monitor displays.

The mean difference (PA minus test measurement) for the First Temp was  $-0.06^{\circ}\text{C}$  (SD  $0.58^{\circ}\text{C}$ ) and the mean difference of the Thermoscan was  $-0.13^{\circ}\text{C}$  (SD  $0.39^{\circ}\text{C}$ ); that is to say PA temperatures were lower than ITT measurements. Both thermometers had similar accuracy. However the variability, as measured by the standard deviation of the First Temp tympanic thermometer was found to be statistically greater than the variability (SD) Thermoscan ( $p < 0.01$ ). This would indicate that the Thermoscan had better reproducibility than the First temp. However, this may be due to an outlier in the First Temp group.

Terndrup & Milewski<sup>7</sup> conducted a study in 1990 of 303 children. They enrolled children under 16 years of age (mean age 1.6 years, standard error 0.9 years) who presented to an emergency department (ED) and underwent rectal thermometry

Two models of tympanic thermometer were used: 1) First Temp used in calibration mode;  $0.8^{\circ}\text{C}$  was added to the measurements (considered the offset to provide rectal equivalence). 2) Thermoscan used in equal mode;  $1.2^{\circ}\text{C}$  was added to the measurements for Thermoscan (considered the offset to provide rectal equivalence). Staff taking measurements were trained and practiced with the two tympanic thermometers. The reference standard used the Diatek electronic rectal thermometer in monitor mode.

This study was of moderate quality. Only age and gender are provided to describe study subjects. Patients were only included if they had a rectal temperature taken. This was routinely used for children under five years or when ordered by the treating physician. Reasons for ordering a rectal temperature have not been documented. All temperature measurements were taken by the same person, therefore these could not be independent or blind.

This study found that the mean difference (rectal temperature minus test temperature) of the uncorrected First Temp and Thermoscan were identical, mean  $1.2^{\circ}\text{C}$ , SD  $0.4^{\circ}\text{C}$ . However the corrected mean of the First Temp (mean  $37.6^{\circ}\text{C}$ , SE  $0.05^{\circ}\text{C}$ ) was found to be significantly lower than the mean rectal temperature (mean  $38.0^{\circ}\text{C}$ , SE  $0.05^{\circ}\text{C}$ ; p value not given). There were no significant differences found between the mean rectal temperature and the mean Thermoscan temperature (mean  $38.0^{\circ}\text{C}$ , SE  $0.05^{\circ}\text{C}$ ). It should be remembered that  $0.8^{\circ}\text{C}$  was added to First Temp and  $1.2^{\circ}\text{C}$  was added to Thermoscan.

Sensitivities and specificities were provided for the two models at selected rectal temperatures. The sensitivity of Thermoscan was higher than First Temp. Conversely the specificity of First Temp was higher than Thermoscan. Receiver Operator Characteristic curve analysis found that the Thermoscan had a larger area under the curve than First Temp (actual figures or statistical testing not provided). Given that the article has not provided any statistical testing, we can not conclude that one model is more accurate than the other.

### Child's age

The systematic review and four cohort studies assessed the effect of child's age on the accuracy of tympanic temperature measurement.

The systematic review by Craig et al<sup>1</sup>, detailed above, found no association between child's age and difference (rectal temperature minus tympanic temperature). Data was only available in 15 of the 31 comparisons, n=2611.

A study by Jean-Mary et al<sup>2</sup> published in 2002 (not detailed when the study was conducted) included 198 infants who were three to 36 months old (mean age 1.3 years) and attended an outpatient department for either 'well child' or acute illness visits. Infants were excluded if they had a contraindication to rectal temperature measurement or had known hypothalamic dysfunction.

The type of ITT used was not documented. Mode used was not documented, but the article stated that the thermometer reflects "an approximation of oral temperature. Oral temperature is generally considered to be 1°F less than rectal temperature. To compensate for this difference 1°F (0.556°C) was added". The reference standard was an Ivac digital rectal thermometer. An infrared axillary thermometer was also used, see below for details. Temperature readings were taken sequentially from each patient by a single experienced nurse, beginning with the least invasive (axillary) and ending with the most invasive (rectal).

This study was of moderate quality. A convenience sample was enrolled in the study. Inclusion criteria were not specified, but it seems all patients were included unless they had either of the two exclusion criteria. All temperature measurements were taken by the same person, therefore could not be independent or blind.

Multiple stepwise regression revealed that child's age was not a significant predictor of ITT accuracy (ITT minus rectal temperature). Data were not provided.

Erickson and Woo<sup>3</sup> (detailed above) found that "None of the temperature methods studied appeared to be significantly affected by the child's age". Data were not presented.

Petersen-Smith et al<sup>4</sup> published their study in 1994 (not detailed when the study was conducted). They enrolled 234 children from birth to 36 months in general paediatric practices. Children with abnormal otic or rectal structures were excluded.

Tympanic temperature was assessed using First Temp Genius model 3000A in rectal mode. The reference standard was a glass-mercury rectal thermometer left in situ for a three-minute period.

This study was of moderate quality. The only descriptor of study participants was gender. Both temperature measurements were taken by the same person, so could not be blind or independent.

The study reported that "age seemed to contribute little to the degree of inaccuracy of this device" that is the ITT measurements. Data were not presented.

Selfridge and Shea<sup>5</sup> conducted a study in 1991. They enrolled 102 infants aged three months and younger presenting to ED. Tympanic temperature was assessed using First Temp model 2000A in oral mode. The reference standard was rectal glass-mercury thermometer. Nurses obtained the measurements, no additional training was provided for tympanic or rectal thermometers.

This study was of moderate quality. Participants were eligible for the study if they presented to the ED and did not require immediate resuscitation. The only descriptor of study participants was gender. Both temperature measurements were taken by the same person, so could not be blind or independent.

The study participants were divided into three age groups: one month or less (n=64); two-months-old (n=25) and three-months-old (n=13). The study found the temperature recording had a clinically significant lower sensitivity in the three-month-old age group (66%) than the one month or less group (90%) and the two-month-old age group (100%). The total number of infants in the three-month-old group was small, and as no statistical testing or confidence intervals were reported, we can not report statistical significance. The specificity, positive and negative predictive values of the groups were similar.

### Otitis Media

It is important for temperature recording to accurately reflect core body temperature rather than localised temperature. Otitis media is infection of the middle ear. The concern with measuring tympanic temperature in patients with otitis media (diagnosed or undiagnosed) is that the temperature recorded may be raised due to hyperaemia rather than correctly reflecting generalised hyperthermia. Otitis media can cause generalised hyperthermia if the infection becomes widespread. However, it is important to be able to differentiate between localised heat and systemic hyperthermia. One review and two cohort studies addressed this issue.

The systematic review by Craig et al<sup>1</sup> (detailed above) found the studies provided insufficient data to undertake this sub-analysis.

Petersen-Smith et al<sup>4</sup> (detailed above) found no significant differences based on findings of otitis media (data not presented). The criteria for otitis media were not stated.

Terndrup and Wong<sup>8</sup> conducted a study in 1989. They enrolled 251 children under 16 years (mean age 21.4 months, SD 22 months) who underwent routine or physician-requested rectal thermometry while in the ED.

Tympanic temperature was assessed using bilateral measurement with First Temp thermometer used in rectal mode by nursing staff blinded to the presence of otitis media. The ear tug was used. The reference standard was Diatek digital/electronic rectal thermometer in monitor mode, allowing at least three minutes for equilibration. Otitis media was recorded after the patient was discharged based on the presence of three or more of the following criteria: 1) rapid onset of fever and/or ear pain and/or irritability, 2) bulging or full tympanic membrane, 3) opaque or erythematous tympanic membrane, 4) limited or no mobility of tympanic membrane.

This study was of moderate quality. Patients were eligible for the study if they presented to ED and underwent either routine or physician-requested rectal temperature measurement. The reasons for physician-requested rectal thermometry were not documented. The only descriptors of study participants were age and gender. It was not documented who took the rectal temperature measurements, and if they were blind or independent of the ITT measurements. The diagnosis of otitis media and temperature taking did appear to be blind and independent. Tympanic temperatures of patients who could not be adequately evaluated were eliminated from the analysis (6 of the 251 children). Reasons for not being adequately evaluated were not documented. However, given the small numbers, this is unlikely to affect results.

This study found there was no significant difference in temperature of infected and uninfected ears. The authors suggest that any elevated temperature recorded in ears with otitis media were likely to be due to systemic temperature increases rather than in local increases.

#### Ear tug

The systematic review and one cohort study looked at the effect of use of ear tug on accuracy of tympanic temperature measurement.

The systematic review by Craig et al<sup>1</sup> (detailed above) found the studies provided insufficient data to undertake this sub-analysis.

Erickson and Woo<sup>3</sup> (detailed above) reported the mean difference (test measurement minus bladder measurement) was not statistically significantly affected by use of ear tug for any of the models ( $p=0.72$ ).

#### Cerumen

Cerumen is commonly called earwax and, usually in small amounts, lines the external ear canal. It is unclear if the presence of excessive wax may obscure the tympanic membrane and thus prevent accurate tympanic temperature measurement. One cohort study assessed the presence of cerumen on the accuracy of tympanic temperature measurement.

Petersen-Smith et al<sup>4</sup> (detailed above) found no significant difference based on the presence of cerumen (figures not presented). How the presence of cerumen was defined was not detailed.

#### Nurse

The technique, training or experience of the person taking the ITT measurements may affect the accuracy of the ITT measurement. One study assessed if the person (nurse) taking the temperature measurement affected the accuracy of tympanic temperature measurement.

Erickson and Wool<sup>3</sup> (detailed above) found "None of the temperature methods studied appeared to be significantly affected by which research nurse made the measurements". Data were not presented. Given that factors such as technique, training or experience of the nurse were not documented, this finding provides us with limited useful information.

## **2. Factors which affect the accuracy of digital/electronic or infrared axillary thermometry**

There were a total of 130 references found with the above search terms. After reviewing the title, abstract or full text, four cohort studies<sup>2,9,10,13</sup> were found which addressed the research question. Two cohort studies were excluded because they undertook insufficient or inappropriate data analysis<sup>10,13</sup>.

Of the studies which assessed axillary temperature measurement, one study used digital/electronic axillary thermometry<sup>5</sup>; this study assessed mode of device. One study used infrared axillary thermometry<sup>2</sup> and this study assessed child's age.

Results of the included studies are summarised here. Full details of appraisals can be found in Appendix 1. The

digital/electronic thermometry results appear first and followed by infrared axillary thermometry results.

#### Mode of device

One cohort study was found which assessed the mode of device and accuracy of axillary temperature measurements and this study only assessed predictive and monitoring modes. No studies were found which assessed the modes which relate to different temperature recording sites in the body. For example core mode and oral mode.

Weiss and Richards<sup>9</sup> conducted a study which was published in 1994 (not detailed when the study was conducted). They enrolled term neonates who were at least 38 weeks gestational age, weighing at least 2350g and in open cribs.

Axillary temperature assessed using IVAC 2080 thermometer in left axilla in both predictive and monitoring modes. The reference standard was an IVAC 2080 in predictive mode inserted at least ¼ inch into the rectum. There were 50 episodes of temperature measurements taken. An episode consisted of three temperatures taken in random order.

This study was of low quality. The only descriptors of study participants were gestational age and birth weight. It is not clear how the participants were chosen. The article has documented that 142 participants were enrolled and that there were 200 episodes of temperature recording in total, 50 episodes in each of the four groups, (we are only including the term neonates in open cribs group). Up to three sets of recordings taken at least one day apart were permitted for each neonate. Taking multiple sets of measurements on individual children may falsely reduce the variability, which may lead to a reduction in the generalisability of the results. It has not been documented how many participants were in each group. All temperature measurements were taken by the same person, so these could not be blind or independent.

The study found there was no significant differences between the thermometer used in predictive mode or in monitoring mode.

#### Child's age

One study assessed the effect of child's age on accuracy of axillary temperature recording; this study assessed infrared axillary thermometry.

Jean-Mary et al<sup>2</sup> (detailed above) used an infrared axillary thermometer, type not specified. Mode used was not documented, but the article stated that the thermometer reflects "an approximation of oral temperature. Oral temperature is generally considered to be 1°F less than rectal temperature. To compensate for this difference 1°F (0.556°C) was added".

Multiple stepwise regression revealed that the accuracy of axillary temperature was significantly, but weakly correlated with child's age (partial  $R^2 = 0.05$ ,  $p < 0.01$ ); the clinical importance of this is minimal.

## Discussion

The studies included in this review were of variable quality with all having some methodological issues. Six studies did not provide specific inclusion and exclusion criteria<sup>2-5,8,9</sup>, and some studies provided limited descriptors of the study participants<sup>2,4,9</sup>. The problem with these issues is that we are unable to know which population the results would be relevant and generalisable to; and if the results are generalisable to the population we are interested in. When undertaking tests of diagnostic accuracy one of the factors which reduces the bias or systematic error is that the outcomes are assessed blindly and independently. All the studies included in this review did not do this as only one assessor recorded the measurements. Although this is a flaw of study design it is difficult to know how this would affect the results.

A major limitation of our review was that many of the studies we found that met the selection criteria were undertaken some time ago. Changes in the technology may reduce the relevance of the results. The reported accuracy of thermometers 10 or so years ago may not apply to current thermometers. The lack of recent studies may also be because we excluded studies which did use core body temperature as the reference standard. Rectal thermometry, for various reasons, has become unpopular in recent times<sup>16</sup>. Therefore this site of core temperature is not used as frequently<sup>16</sup>.

### **1. Infrared tympanic thermometry**

#### Mode of device

There are recognised variations between different temperature recording sites in the body. The manufacturers of ITT have allowed for this by adding an offset to the temperature recording of their machines. Although the systematic review<sup>1</sup> was unable to report any statistically significant differences between the modes used there does appear to be clinically relevant trends. That is to say a tympanic thermometer set on tympanic or actual mode tended to report a lower temperature than one set on core mode.

The emergency department at Casey Hospital use their tympanic thermometers on oral mode. The systematic review<sup>1</sup> found this mode had a pooled mean difference of 0.34°C less than core (rectal) temperature. That is to say, on average the device will record a temperature measurement 0.34°C less than the rectal temperature measurement.

### Model of device

The three studies<sup>3,5,7</sup> which looked at model of device reported no significant difference in accuracy between Thermoscan and First Temp thermometers. One study did report that the variability of the First Temp was more than that of the Thermoscan<sup>5</sup>; however, there was an outlier in the data which may account for this. One study assessed Core-Check in addition to First Temp and Thermoscan<sup>3</sup> and found this device to be statistically less accurate than the latter two devices.

The emergency department at Casey Hospital, use the Genius 3000A. This machine was used in one study<sup>3</sup> but the other studies<sup>5,7</sup> did not document the type of model used.

### Child's age

Age was not found to be a factor affecting the accuracy of tympanic temperature measurement. It should be noted that one study found temperature measurement had a clinically significant lower sensitivity in the three-month-old age group than the one month or less group and the two-month-old age group<sup>5</sup>. The total number of infants in the three-month-old group was small, and no statistical testing or confidence intervals were reported.

### Otitis Media

Three studies looked at the effect of otitis media on the accuracy of tympanic temperature measurement<sup>1,4,8</sup>. The systematic review<sup>1</sup> found insufficient data to undertake this analysis. The other two cohort studies<sup>4,8</sup> reported no significant effect of otitis media; only one of these studies reported clear criteria for this condition<sup>8</sup>.

### Ear tug

Only two studies assessed the effect of the ear tug on accuracy of tympanic temperature recording<sup>1,3</sup>. The systematic review<sup>1</sup> found insufficient data to undertake this analysis. The other study reported no significant effect of the ear tug on the accuracy of tympanic temperature measurement<sup>3</sup>.

### Cerumen

Only one study looked at this factor<sup>4</sup>, but did not document how the presence of cerumen was defined. Therefore the generalisability of their findings is limited.

### Nurse

Only one study assessed if the nurse measuring the temperature affected the accuracy of tympanic temperature measurement<sup>3</sup>. Factors such as technique, training or experience of the nurse would have been important to know. However as this information was not provided, the results are of limited value.

## **2. Digital/electronic or infrared axillary thermometry**

### Mode of device

The emergency department at Casey Hospital use their axillary thermometers in predictive mode, Only one study assessed if predictive or monitoring mode affected the accuracy of axillary temperature measurement<sup>9</sup>. This study reported that there was no difference between the two modes.

### Child's age

The one study that assessed the effect of child's age on the accuracy of axillary temperature measurement reported no clinical significant effect<sup>2</sup>.

## **Conclusions**

### Infrared tympanic thermometry

- Although there were no statistically significant differences found between the modes used, there were some clinically important differences. The smallest differences were seen when the using the ITT in rectal and core modes. That is to say these modes provided the most accurate result. When actual and tympanic modes were used there were the largest differences; these modes provided the least accurate result.
- Model of device, child's age, presence of otitis media and the ear tug do not appear to affect the accuracy of the tympanic temperature measurements.
- There is insufficient evidence to draw any conclusions about how the presence of cerumen or the nurse taking the measurements affects the accuracy of infrared tympanic thermometer measurement.

### Digital/electronic axillary thermometry

- There is insufficient evidence to draw any conclusions about how the mode of device (predictive or monitoring) affects the accuracy of digital/electronic axillary temperature measurements

### Infrared axillary thermometry

- There is insufficient evidence to draw any conclusions about how child's age affects the accuracy of infrared axillary temperature measurements.

## References

1. Craig, J.V., et al., *Infrared ear thermometry compared with rectal thermometry in children: a systematic review.[see comment]*. *Lancet*, 2002. **360**(9333): p. 603-9.
2. Jean-Mary, M.B., et al., *Limited accuracy and reliability of infrared axillary and aural thermometers in a pediatric outpatient population*. *Journal of Pediatrics*, 2002. **141**(5): p. 671-6.
3. Erickson, R.S. and T.M. Woo, *Accuracy of infrared ear thermometry and traditional temperature methods in young children*. *Heart & Lung*, 1994. **23**(3): p. 181-95.
4. Petersen-Smith, A., et al., *Comparison of aural infrared with traditional rectal temperatures in children from birth to age three years*. *The Journal of Pediatrics*, 1994. **125**(1): p. 83-85.
5. Selfridge, J. and S.S. Shea, *The accuracy of the tympanic membrane thermometer in detecting fever in infants aged 3 months and younger in the emergency department setting*. *Journal of Emergency Nursing*, 1993. **19**(2): p. 127-30.
6. Romano, M.J., et al., *Infrared tympanic thermometry in the pediatric intensive care unit*. *Critical Care Medicine*, 1993. **21**(8): p. 1181-5.
7. Terndrup, T.E. and A. Milewski, *The performance of two tympanic thermometers in a pediatric emergency department*. *Clinical Pediatrics*, 1991. **30**(4 Suppl): p. 18-23; discussion 34-5.
8. Terndrup, T.E. and A. Wong, *Influence of otitis media on the correlation between rectal and auditory canal temperatures*. *American Journal of Diseases of Children*, 1991. **145**(1): p. 75-8.
9. Weiss, M.E. and M.T. Richards, *Accuracy of electronic axillary temperature measurement in term and preterm neonates*. *Neonatal Network - Journal of Neonatal Nursing*, 1994. **13**(8): p. 35-40.
10. Wilshaw, R., et al., *A comparison of the use of tympanic, axillary, and rectal thermometers in infants*. *Journal of Pediatric Nursing*, 1999. **14**(2): p. 88-93.
11. Yetman, R.J., et al., *Comparison of temperature measurements by an aural infrared thermometer with measurements by traditional rectal and axillary techniques*. *Journal of Pediatrics*, 1993. **122**(5 Pt 1): p. 769-73.
12. Stewart, J.V. and D. Webster, *Re-evaluation of the tympanic thermometer in the emergency department.*. *Annals of Emergency Medicine*, 1992. **21**(2): p. 158-61.
13. Muma, B.K., et al., *Comparison of rectal, axillary, and tympanic membrane temperatures in infants and young children*. *Annals of Emergency Medicine*, 1991. **20**(1): p. 41-4.
14. Altman, D.G. and J.M. Bland, *Diagnostic tests. 1: Sensitivity and specificity*. *BMJ*, 1994. **308**(6943): p. 1552.
15. Altman, D.G. and J.M. Bland, *Diagnostic tests 3: receiver operating characteristic plots*. *BMJ*, 1994. **309**(6948): p. 188.
16. Dew, P.L., *Is tympanic membrane thermometry the best method for recording temperature in children?* *Journal of Child Health Care*, 2006. **10**(2): p. 96-110.

## Disclaimer

The information in this report is a summary of that available and is primarily designed to give readers a starting point to consider currently available research evidence. Whilst appreciable care has been taken in the preparation of the materials included in this publication, the authors and Southern Health do not warrant the accuracy of this document and deny any representation, implied or expressed, concerning the efficacy, appropriateness or suitability of any treatment or product. In view of the possibility of human error or advances of medical knowledge the authors and Southern Health cannot and do not warrant that the information contained in these pages is in every aspect accurate or complete. Accordingly, they are not and will not be held responsible or liable for any errors of omissions that may be found in this publication. You are therefore encouraged to consult other sources in order to confirm the information contained in this publication and, in the event that medical treatment is required, to take professional expert advice from a legally qualified and appropriately experienced medical practitioner.

## Appendix 1: Critical Appraisal of Identified Research

### 1. Infrared tympanic thermometers

#### Appraisal of included systematic review:

**Study:** Craig et al<sup>1</sup>, 2002

#### Description of study – systematic review

<b>Patients</b>	Children 0 to 18 years.
<b>N</b>	4441 included (25 studies, 31 comparisons), data reported on only 4344
<b>When conducted</b>	Search conducted Dec 1999 and in 2000 - exact date in 2000 not specified
<b>Setting</b>	Various
<b>Infrared tympanic</b>	
1. <b>Type</b>	1. Various – not always documented in the studies
2. <b>Mode used</b>	2. Various – not always documented in the studies
3. <b>Ear tug</b>	3. Various – not always documented in the studies
4. <b>Otitis media</b>	4. Various – not always documented in the studies
5. <b>Calibration</b>	5. Various – not always documented in the studies
6. <b>Technique</b>	6. Various – not always documented in the studies
<b>Comparison factors</b>	Mode in which temperature is taken (rectal, actual, core, oral, tympanic, not stated). Other factors: child's age, otitis media and ear tug.
<b>Outcomes</b>	Temperature
<b>Reference standard</b>	
1. <b>Type</b>	1. Rectal temperature assessed using electronic, glass-mercury or indwelling thermocouple device
2. <b>Mode used</b>	2. Not specified
3. <b>Calibration</b>	3. Not specified
4. <b>Technique</b>	4. Not specified
<b>Inclusion Criteria</b>	Included method comparison studies in which temperature taken at the ear (test site) was compared to rectal temperature (reference site).
<b>Exclusion criteria</b>	Excluded with hypothermia (rectal temp <35.0°C) or preterm (born at <37 weeks' gestation).

## Appraisal of included cohort studies:

**Study:** Jean-Mary et al<sup>2</sup>, 2002

### Description of study

<b>Patients</b>	Children 3–36 months old
<b>N</b>	198
<b>When conducted</b>	Not detailed when the study was conducted
<b>Setting</b>	An outpatient department for either 'well child' or acute illness visits in the US
<b>Infrared tympanic</b>	
1. <b>Type</b>	1. Type not specified.
2. <b>Mode used</b>	2. Mode used not documented, but the manufacture's handbook states that the temperature reflects an approximation of oral temperature.
3. <b>Ear tug</b>	3. Ear tug not specified
4. <b>Otitis media</b>	4. Otitis media not specified
5. <b>Calibration</b>	5. Calibrated according to hospital protocol or manual
6. <b>Technique</b>	6. Used as described in the user manual. 1°F (0.556°C) added to bring in line with rectal temperature.
<b>Comparison factors</b>	Child's age
<b>Outcomes</b>	Three temperature readings were taken sequentially from each patient by a single experienced nurse, beginning with the least invasive (axillary) and ending with the most invasive (rectal) Fever defined as rectal temperature ≥ 100.4°F (38.0°C)
<b>Reference standard</b>	
1. <b>Type</b>	1. An Ivac digital rectal thermometer
2. <b>Mode used</b>	2. Mode used not specified
3. <b>Calibration</b>	3. Calibrated according to hospital protocol or manual
4. <b>Technique</b>	4. Technique not specified
<b>Inclusion Criteria</b>	Nil other than specified above
<b>Exclusion criteria</b>	Contraindication to rectal temperature measurement or known hypothalamic dysfunction

**Study:** Erickson and Wool<sup>3</sup>, 1994

**Description of study**

<b>Patients</b>	One to 45 months		
<b>N</b>	30		
<b>When conducted</b>	Study conducted between March and August 1992		
<b>Setting</b>	Intensive care and cardiac recovery units in the US		
<b>Infrared tympanic</b>			
<b>1. Type</b>	1. Core-check model 2090A	1. First Temp Genius model 3000A	1. Thermoscan Pro-1 model IR-1
<b>2. Mode used</b>	2. Used in core mode		
<b>3. Ear tug</b>	3. Two readings were made with each thermometer, one using ear tug and the other without the ear tug.		
<b>4. Otitis media</b>	4. Otitis media not specified		
<b>5. Calibration</b>	5. Calibrated before, during and after data collection in the unadjusted mode		
<b>6. Technique</b>	6. All six readings were made in the same ear in random order, with at least one minute apart to allow recovery from transient tissue cooling due to contact with the comparatively cool probe.		
<b>Comparison factors</b>	Model of ITT, use of ear tug, child's age or research nurse taking the measurements		
<b>Outcomes</b>	Bladder, rectal and axillary temperatures were read from continuous digital displays immediately after each ear-based measurement. Axillary readings were made three minutes before and two minutes after the six tympanic measurements were taken, for a total of 11 minutes.  Fever defined as bladder temperature $\geq 38^{\circ}\text{C}$ , model A $\geq 37.3^{\circ}\text{C}$ , model B $\geq 37.7^{\circ}\text{C}$ , model C $\geq 37.5^{\circ}\text{C}$		
<b>Reference standard</b>			
<b>1. Type</b>	1. Bladder temperatures were continuously monitored using: Mon-a-therm Temperature Monitoring System model 4070 (thermistor) or Mon-a-therm CathTemp used with Mon-a-therm Temperature Monitoring System model 6510 (thermocouple)		
<b>2. Mode used</b>	2. Mode used not specified		
<b>3. Calibration</b>	3. The sensor-monitor systems were tested in the "water bath"		
<b>4. Technique</b>	4. Measurements read from a continuous digital display		
	Also used a rectal (thermocouple) to record rectal temperatures.		

<b>Inclusion Criteria</b>	Children who required bladder catheters for their care.
<b>Exclusion criteria</b>	Nil documented
<b>Study:</b> Petersen-Smith et al <sup>4</sup> , 1994	
<b>Description of study</b>	
<b>Patients</b>	Children from birth to 36 months in general paediatric practices
<b>N</b>	234
<b>When conducted</b>	Not documented
<b>Setting</b>	US
<b>Infrared tympanic</b>	
<b>1. Type</b>	1. First Temp Genius model 3000A
<b>2. Mode used</b>	2. Used in rectal mode
<b>3. Ear tug</b>	3. Not specified
<b>4. Otitis media</b>	4. Presence of otitis media was recorded
<b>5. Calibration</b>	5. Calibrated before the study.
<b>6. Technique</b>	6. Investigators received, from the manufacture, written and oral instruction. Temperatures were taken within five minutes of initiation or completion of the rectal measurement.
<b>Comparison factors</b>	Child's age, otitis media and cerumen
<b>Outcomes</b>	Temperature
<b>Reference standard</b>	
<b>1. Type</b>	1. Glass-mercury rectal thermometer,
<b>2. Mode used</b>	2. Not applicable
<b>3. Calibration</b>	3. Calibrated before the study.
<b>4. Technique</b>	4. Measured for a 3-minute period.
<b>Inclusion Criteria</b>	As above
<b>Exclusion criteria</b>	Children having abnormal otic or rectal structures were excluded.

**Study:** Selfridge and Shea<sup>5</sup> 1993**Description of study**

<b>Patients</b>	Infants aged 3 months and younger
<b>N</b>	102
<b>When conducted</b>	Study conducted from Jan to July 1991
<b>Setting</b>	Emergency department of a US hospital
<b>Infrared tympanic</b>	
1. <b>Type</b>	1. First Temp model 2000A
2. <b>Mode used</b>	2. Used in oral mode
3. <b>Ear tug</b>	3. Not specified
4. <b>Otitis media</b>	4. Not specified
5. <b>Calibration</b>	5. Calibrated by the hospital biomedical engineering department to comply with accuracy standards before the study inception
6. <b>Technique</b>	6. Recorded in degrees Fahrenheit. ED nurses obtained the measurements, no additional training was provided
<b>Comparison factors</b>	Child's age
<b>Outcomes</b>	Temperatures were recorded sequentially, the tympanic and then rectal measurement Fever was defined as temp >99.6°F (37.6°C) by ITT or temp > 100.6°F (38.1°C) by rectal thermometer One degree of temperature difference was used as equipment literature from the manufacturer supports an approximate 1° F difference between rectal and oral readings (0.8°F ±0.2°F).
<b>Reference standard</b>	
1. <b>Type</b>	1. Rectal glass-mercury thermometer
2. <b>Mode used</b>	2. Not applicable
3. <b>Calibration</b>	3. Not specified
4. <b>Technique</b>	4. ED nurses obtained the measurements, no additional training was provided
<b>Inclusion Criteria</b>	As above
<b>Exclusion criteria</b>	Not requiring immediate resuscitation

**Study:** Romano et al<sup>6</sup>, 1993

**Description of study**

<b>Patients</b>	Six months to 15 years	
<b>N</b>	20	
<b>When conducted</b>	Study conducted between July 1991 and March 1992.	
<b>Setting</b>	ICU of a tertiary paediatric hospital in the US	
<b>Infrared tympanic</b>		
<b>1. Type</b>	1. Genius First Temp	1. Thermoscan Pro-1
<b>2. Mode used</b>	2. Used in core mode	2. Used on core mode
<b>3. Ear tug</b>	3. No ear tug was used	3. The ear tug was used as recommended by the manufacturer
<b>4. Otitis media</b>	4. Otitis media not specified	4. Otitis media not specified
<b>5. Calibration</b>	5. Calibrated using a manufacturer-supplied “black-body” temperature reference source before patient enrolment and at the midpoint of the study	5. Calibrated using a manufacturer-supplied “black-body” temperature reference source before patient enrolment and at the midpoint of the study.
<b>6. Technique</b>	6. Measurements were taken by one of the authors, who were instructed by a manufacturer’s representative on the proper use	6. Measurements were taken by one of the authors, who were instructed by a manufacturer’s representative on the proper use
<b>Comparison factors</b>	Model of ITT	
<b>Outcomes</b>	All measurements were taken ‘simultaneously’ by one of the authors	
<b>Reference standard</b>		
<b>1. Type</b>	1. Opticath pulmonary artery catheters were used	
<b>2. Mode used</b>	2. Mode not applicable	
<b>3. Calibration</b>	3. Calibration not specified	
<b>4. Technique</b>	4. Inserted using standard techniques	
	Also used a rectal probe (Mon-A-Therm) which was compared to the pulmonary artery temperature to allow comparison of variability.	
<b>Inclusion Criteria</b>	Admitted to paediatric ICU and requiring pulmonary artery catheter	
<b>Exclusion criteria</b>	Inability to visualise tympanic membrane, presence of tympanostomy tubes or contraindication to the use of indwelling rectal tubes.	

**Study:** Terndrup and Milewski<sup>7</sup>, 1991

**Description of study**

<b>Patients</b>	Children under 16 years of age presenting to ED	
<b>N</b>	303	
<b>When conducted</b>	Study conducted between Aug and Oct 1990	
<b>Setting</b>	Emergency department of a US hospital	
<b>Infrared tympanic</b>		
<b>1. Type</b>	1. First Temp	1. Thermoscan
<b>2. Mode used</b>	2. Used in calibration mode	2. Used in equal mode
<b>3. Ear tug</b>	3. Ear tug was used	3. Ear tug was not used
<b>4. Otitis media</b>	4. Not stated	4. Not stated
<b>5. Calibration</b>	5. Thermometers are self-calibrating, but were validated against a black body.	5. Thermometers are self-calibrating, but were validated against a black body
<b>6. Technique</b>	6. 0.8°C was added to the measurements for First Temp (standard offset for rectal equivalence for First Temp).	6. 1.2°C was added to the measurements for Thermoscan (standard offset for rectal equivalence for Thermoscan)
<b>Comparison factors</b>	Model of ITT	
<b>Outcomes</b>	Staff taking measurements were trained and practiced with the two tympanic thermometers.	
<b>Reference standard</b>		
<b>1. Type</b>	1. Diatek digital/electronic rectal thermometer	
<b>2. Mode used</b>	2. Used in monitor mode	
<b>3. Calibration</b>	3. Calibrated against a water bath at the beginning and at the end of the study.	
<b>4. Technique</b>	4. Allowed at least 3 minutes for equilibration	
<b>Inclusion Criteria</b>	As above and underwent rectal thermometry	
<b>Exclusion criteria</b>	Recent auditory surgery, haemorrhage or suspected cerebrospinal fluid extravasating for the ear	

**Study:** Terndrup and Wong<sup>8</sup>, 1991

**Description of study**

<b>Patients</b>	Children under 16 years
<b>N</b>	251 in US
<b>When conducted</b>	Study conducted in Nov and Dec 1989.
<b>Setting</b>	US
<b>Infrared tympanic</b>	
1. <b>Type</b>	1. First Temp thermometer
2. <b>Mode used</b>	2. Used in rectal mode
3. <b>Ear tug</b>	3. Ear tug used
4. <b>Otitis media</b>	4. Patients were assessed for this—see below
5. <b>Calibration</b>	5. The probe and microprocessor have an internal, self-adjusting calibration system mode
6. <b>Technique</b>	6. Bilateral measurements were taken
<b>Comparison factors</b>	Otitis media - recorded following pt discharge, based on the presence of three or more of the following criteria: <ol style="list-style-type: none"><li>1. rapid onset of fever and/or ear pain and/or irritability</li><li>2. bulging or full tympanic membrane</li><li>3. opaque or erythematous tympanic membrane</li><li>4. limited or no mobility of tympanic membrane</li></ol>
<b>Outcomes</b>	Temperature recording
<b>Reference standard</b>	
1. <b>Type</b>	1. Diatek digital/electronic rectal thermometer
2. <b>Mode used</b>	2. Used in monitor mode
3. <b>Calibration</b>	3. Temperature calibrations were performed against a water bath at the beginning and end of the experimental period.
4. <b>Technique</b>	4. Allowed at least 3 minutes for equilibration
<b>Inclusion Criteria</b>	As above plus they underwent either routine or physician-requested rectal thermometry
<b>Exclusion criteria</b>	Nil stated

## 2. Digital/electronic or infrared axillary thermometers

Appraisal of included cohort studies	
<b>Study:</b> Jean-Mary et al <sup>2</sup> , 2002	
<b>Description of study</b>	
<b>Patients</b>	Children 3-36 months old
<b>N</b>	198
<b>When conducted</b>	Not detailed when the study was conducted
<b>Setting</b>	An outpatient department for either 'well child' or acute illness visits in the US
<b>Axillary</b>	
1. <b>Type</b>	1. Infrared thermometer, type not specified
2. <b>Mode used</b>	2. Mode used not documented, but the manufacture's handbook states that the temperature reflects an approximation of oral temperature.
3. <b>Calibration</b>	3. Calibrated according to hospital protocol or manual
4. <b>Technique</b>	4. Used as described in the user manual 1°F (0.556°C) added to bring in line with rectal temperature.
<b>Comparison factors</b>	Child's age
<b>Outcomes</b>	Three temperature readings were taken sequentially from each patient by a single experienced nurse, beginning with the least invasive (axillary) and ending with the most invasive (rectal). Fever defined as rectal temperature $\geq 100.4^{\circ}\text{F}$ ( $38.0^{\circ}\text{C}$ ).
<b>Reference standard</b>	
1. <b>Type</b>	1. An Ivac digital rectal thermometer
2. <b>Mode used</b>	2. Mode used not specified
3. <b>Calibration</b>	3. Calibrated according to hospital protocol or manual
4. <b>Technique</b>	4. Technique not specified
<b>Inclusion Criteria</b>	Nil other than specified above
<b>Exclusion criteria</b>	Contraindication to rectal temperature measurement or known hypothalamic dysfunction

---

**Study:** Weiss and Richards<sup>9</sup> 1994

---

**Description of study**

---

<b>Patients</b>	Term neonates, up to three months old
<b>N</b>	Number not provided. There were 50 episodes of temperature recording.
<b>When conducted</b>	Not documented when the study was undertaken
<b>Setting</b>	In US.
<b>Axillary</b>	
1. <b>Type</b>	1. IVAC 2080
2. <b>Mode used</b>	1. A. Predictive or B. monitoring
3. <b>Calibration</b>	2. Checked for calibration by the hospital's biomedical engineering department
4. <b>Technique</b>	3. Taken in the left axilla
<b>Comparison factors</b>	Mode of device: (Predictive or monitoring mode)
<b>Outcomes</b>	<p>The nurse caring for the neonate took the measurements. The research assistant monitored the nurse's temperature taking technique and gave direction on corrected placement, as needed.</p> <p>Up to three recording episodes per infant, at least one day apart. An episode consisted of three temperatures taken in random order.</p>
<b>Reference standard</b>	
1. <b>Type</b>	1. IVAC 2080 rectal thermometer
2. <b>Mode used</b>	2. Used in predictive mode,
3. <b>Calibration</b>	3. Checked for calibration by the hospital's biomedical engineering department.
4. <b>Technique</b>	4. Inserted the rectal probe at least ¾ inch into the rectum.
<b>Inclusion Criteria</b>	At least 38 weeks gestational age, weighing at least 2500 gms, in open cribs
<b>Exclusion criteria</b>	Nil specified

---

Quality of included systematic review									
Study	Focused research question	Specified inclusion/exclusion criteria	Explicit and comprehensive search strategy	Reviewers blind to author, institution & affiliations	Validity of included trials appraised	Homogeneity between studies assessed	Summary of main results presented	Strengths and limitations of included studies discussed	Comments
Craig et al <sup>1</sup>	Yes	Yes	Yes	Not stated	Yes	Yes	Yes	Yes	Validity of trials appraised, but the assessors used limited criteria and did not discuss the strengths and limitations of the studies in great detail.
Quality of included cohort studies									
Study	Specified inclusion/exclusion criteria	Explicit description of study subjects	Appropriate spectrum of consecutive patients who would normally be tested for the disorder of interest and who disease status is not known.	Use of appropriate 'gold standard' reference test	All participants are assessed with both study test and reference standard test	Assessment of test outcomes are independent	Assessors are blind to result of other test	Both sensitivity and specificity, or number of true positives, false positives, true negatives and false negatives	Comments
Jean-Mary et al <sup>2</sup>	Some	No	Unknown	Yes	Yes	No	No	Yes	Inclusion criteria not specified, but it seems all patients were included, unless they had either of the two exclusion criteria (contraindication to rectal temperature measurement or those with known hypothalamic dysfunction). A convenience sample was enrolled in the study; it has not been documented how these were chosen. All temperature measurements were taken by the same person, therefore these could not be independent or blind. 1°F (0.56°C) was added to the ITT and axillary temp as 'Oral temperatures are generally considered to be 1°F less than rectal temperature.'
Erickson and Woo <sup>3</sup>	No	Yes	No	Yes	Yes	No	No	Yes	The study subjects were a convenience sample who required bladder catheter; no other inclusion/exclusion criteria documented. Each participant had six ear measurements taken, and for each of these there was a corresponding axillary, bladder and rectal temperature taken – (6 sets of measurements, 24 on each child (720 total measurements). ). Taking multiple sets of measurements on individual children may falsely reduce the variability, which may lead to a reduction in the generalisability of the results. All temperature measurements were taken by the same person, therefore

									these could not be independent or blind. It appears these were objective. No power calculations appear to have been undertaken, so the study may have been underpowered.
Petersen-Smith et al <sup>4</sup>	Some	Some	Yes	Yes	Yes	No	No	No	Children were excluded if they had abnormal otic or rectal structures – not clear if there were other inclusion/exclusion criteria. The only descriptor of study participants was gender. Both temperature measurements were taken by the same person, so could not be blind or independent.
Selfridge and Shea <sup>5</sup>	No	No	Yes	Yes	Yes	No	No	Yes	Patients were eligible for the study if they presented to ED and didn't require immediate resuscitation. The only descriptor of study participants was gender. Both temperature measurements were taken by the same person, so could not be blind or independent.
Romano et al <sup>6</sup>	Yes	Some	Yes	Yes	Yes	No	No	No	Little information provided about the participants. All temperature measurements were taken by the same person, therefore these could not be independent or blind. It appears they were objective. Sensitivity or specificity not provided, but rather information about mean bias and standard deviation. There are outliers in the data for the First Temp thermometer, this would have increased the variability of this method and may be the reason why this method has statistically greater variability than the PA and rectal temperature. It would have been useful to analyse the data with and without the outliers, to see if this affected the result.
Terndrup and Milewski <sup>7</sup>	Yes	Some	Unclear	Yes	Yes	No	No	Yes	Only age and gender are provided to describe study subjects. Patients were only included if they had a rectal temperature. This was routine for children under 5 years, and when ordered by the treating physician. Reasons for ordering a rectal temperature have not been documented. All temperature measurements were taken by the same person, therefore these could not be independent or blind. Staff were discouraged from repeating possibly spurious measurements unless error messages occurred.
Terndrup and Wong <sup>8</sup>	No	Some	Unclear	Yes	Yes	No	No	Yes	Patients were eligible for the study if they presented to ED and a rectal measurement was requested. Reasons for rectal thermometry not documented. Only descriptors of study participants are age and gender. Not clear who took the temperature measurements, but it seems unlikely these were blind or independent of each other. The diagnosis of otitis media and temperature taking appear blind and independent. Tympanic temperatures of patient who could not be adequately evaluated were eliminated from the analysis (6 of 251). Reasons for not being adequately evaluated were not documented, and given the small numbers this is unlikely to affect results.

Weiss and Richards <sup>9</sup>	No	No	Unclear	Yes	Yes	No	No	No	<p>The only descriptor of study participants are gestational age and birth weight.</p> <p>Not clear how the participants were chosen.</p> <p>The article has documented that 142 pts were enrolled and that there were 200 episodes of temperature recording in total, 50 in each of the 4 groups, (we are only appraising one group – term neonates in open cribs). It has not been documented how many participants were in each group.</p> <p>Participants could have up to 3 measurements each (at least a day apart). Taking multiple sets of measurements on individual children may falsely reduce the variability, which may lead to a reduction in the generalisability of the results.</p> <p>Both temperature measurements were taken by the same person, so could not be blind or independent.</p>
---------------------------------	----	----	---------	-----	-----	----	----	----	--

## Results of included studies:

**Craig et al<sup>1</sup>:** (ITT) (quality of study – high)

### MODEL OF DEVICE

Random effects pooled mean was:

- Rectal mode 0.15°C (SD 0.55°C);
- Actual mode 0.70°C (SD 0.45°C);
- Core mode 0.25°C (SD 0.51°C);
- Oral mode 0.34°C (SD 0.60°C);
- Tympanic mode 0.62°C (SD 0.51°C);
- Mode not stated 0.32°C (SD 0.45°C).
- There was significant residual heterogeneity between mean biases within mode groups ( $p < 0.0001$  for all modes). That is to say the results of the studies were too different to be able to be certain as to what the true mean bias is.
- There was significant residual heterogeneity between SDs within mode groups (rectal mode  $p < 0.0001$ ; actual mode  $p = 0.0111$ , core mode  $p < 0.0001$ ; oral mode  $p = 0.0339$ ; mode not stated  $p < 0.0001$ ).
- Although there does not appear to be any statistically significant difference between the modes of the devices, there does seem to be clinically significant differences between the modes: rectal and core modes having the least bias; and actual and tympanic modes having the most bias.

### CHILD'S AGE

No association was found between age and bias between rectal and tympanic temperature (data only available in 15 of the 31 comparisons,  $n = 2611$ ).

### OTITIS MEDIA

Insufficient data provided in the studies for the authors (of the systematic review) to undertake this sub-analysis.

### EAR TUG

Insufficient data provided in the studies for the authors (of the systematic review) to undertake this sub-analysis.

---

**Jean-Mary et al<sup>2</sup>** (quality of study – moderate)

### Infrared axillary thermometer:

#### CHILD'S AGE

Multiple stepwise regression revealed that the bias of axillary temperature was significantly, but very weakly correlated with child's age (partial  $R^2 = 0.05$ ,  $p < 0.01$ ); the clinical significance of this is minimal. The direction of this association is not detailed, but from the information provided on age groups sub-analysis, it would appear that the bias tended to be more negative in the 1-3 year old group than the  $< 1$  year group.

### Infrared tympanic thermometer:

#### CHILD'S AGE

Multiple stepwise regression revealed that age was not a significant predictor of aural temperature bias.

---

**Erickson and Woo<sup>3</sup>** (quality of study – moderate)

### Tympanic thermometer

#### MODEL

Tested 3 type of thermometers;

- The mean biases (test measurement – bladder measurement) of the different models were: model A  $-0.68^\circ\text{C}$  (SD  $0.41^\circ\text{C}$ ), model B  $-0.29^\circ\text{C}$  (SD  $0.41^\circ\text{C}$ ) and model C  $-0.46^\circ\text{C}$  (SD  $0.50^\circ\text{C}$ ). Analysis of variance showed that these differences were statistically significantly ( $p < 0.01$ ) and post-hoc testing indicated that model A's values were significantly more negative (less close to bladder temperature) than those with model B ( $p < 0.01$ ) and also significantly more negative than those of model C's ( $p = 0.04$ ). That is to say Model B was the most accurate, then model C; model A was the least accurate.
- The variance in the measurements taken by the different models was similar with standard deviations of  $0.4^\circ\text{C}$  to  $0.5^\circ\text{C}$ . This variation is smaller than in other studies, this is likely to be because 6 sets (24 actual measurements) were taken for each child.

- Sensitivity of the tympanic thermometers to detect fever ranged from 65% to 92%; specificity from 81% to 91%. (Model A 65% and 91%, Model B 75% and 81%, Model C 92% and 86%, sensitivity and specificity, respectively). NB: fever defined as temp at recorded at the bladder  $\geq 38^{\circ}\text{C}$ , model A  $37.3^{\circ}\text{C}$ , model B  $37.7^{\circ}\text{C}$ , model C  $37.5^{\circ}\text{C}$ . The different criteria for fever are due to the mean bias of the different devices. Model C does appear clinically better able to differentiate between fever and non-fever than Model A or Model B, but as no statistic testing or CIs have been reported, the statistical significance is unclear.
- Taking all the above information into account, we conclude that Model B (First Temp Genius – Model 3000A) or Model C (THERMOSCAN PRO-1, Model IR-1) were similar in accuracy and Model A (CORE-CHECK, Model 2090A) was the least accurate.

#### EAR TUG

- The mean bias was not statistically significantly affected by use of ear tug for any of the models ( $p=0.72$ ).

#### OTHER FACTORS

- “None of the temperature methods studied appeared to be significantly affected by the child’s age or which research nurse made the measurements”. It is unclear if it is the mean bias (test measurement minus bladder measurement) or actual temperature (by each device) that is being tested. (Data not presented).

#### Axillary

##### PLACEMENT

- After exposing the axilla for probe placement, the area gradually warmed from a mean bias of  $-1.27^{\circ}\text{C}$  at 1 minute, to  $-1.02^{\circ}\text{C}$  at 3 minutes,  $-0.89^{\circ}\text{C}$  at 5 minutes and  $-0.81^{\circ}\text{C}$  at 11 minutes ( $p<0.01$ ) The children reached their maximum temperature in 1 to 11 minutes. Only 13% had done so at 3 minutes, 20% at 5 minutes, 70% at 9 minutes, and 90% at 10 minutes. This would appear to indicate that arm abduction can affect axillary temperature, and is likely to be a clinical important finding.

#### OTHER FACTORS

- None of the temperature methods studied appeared to be significantly affected by the child’s age or which research nurse made the measurements.

#### **Petersen-Smith et al<sup>4</sup>** (quality of study – moderate)

##### CHILD’S AGE

The study found that “age seemed to contribute little to the degree of inaccuracy of this device” ie ITT measurements (data not presented).

##### OTITIS MEDIA

The study found no significant differences based on findings of otitis media (data not presented). The criteria for otitis media were not been stated.

##### CERUMEN

The study found no significant differences based on the presence of cerumen (data not presented). How the presence of cerumen was defined has not been detailed.

#### **Selfridge and Shea<sup>5</sup>** (quality of study – moderate)

##### CHILD’S AGE

The divided the study participants into three groups: aged 1 month or less ( $n=64$ ); 2-months-old ( $n=25$ ) and 3-months-old ( $n=13$ ). They found:

- The 3-month-old age group had a clinically significant lower sensitivity than the 1 month or less group and the 2-month-old age group (66% vs 90% and 100%). The total number of infants in this group was small, which may account for the low sensitivity value. No statistical testing was reported.
- The specificity, positive and negative predictive values of the groups were similar. The subset of patients aged 1 month or less, specificity was 88%, positive predictive value was 60% and the negative predictive value was 98%. In the 2-month-old age group specificity was 90%, positive predictive value was 61% and the negative predictive value was 100%. In the 3-month-old age group ( $n=13$ ) specificity was 90%, the positive predictive value was 66% and the negative predictive value was 90%.
- Overall sensitivity and specificity, of the ITT to detect fever, were 88% and 89%, respectively. The positive predictive value was 74% and the negative predictive value was 97%.

#### **Romano et al<sup>6</sup>** (quality of study – moderate)

### Tympanic thermometer

- The mean bias was calculated as pulmonary artery temperature minus test measure, and the mean bias between PA and ITT (First Temp) was  $-0.06^{\circ}\text{C}$  (SD  $0.58^{\circ}\text{C}$ ); PA and ITT (Thermoscan) was  $-0.13^{\circ}\text{C}$  (SD  $0.39^{\circ}\text{C}$ ) (ie tympanic temperatures were lower). The mean biases of the tympanic thermometers were not significantly different from that of the PA. The variability (SD) of the First Temp tympanic thermometer showed significantly greater variability than the SD of the PA minus rectal temperatures ( $p < 0.01$ ); whereas the Thermoscan model did not.
  - The accuracy (mean bias) was similar for both of the tympanic thermometers. The variability (SD) of the First Temp tympanic thermometer was statistically greater than the variability of the PA and the rectal temperature. However there were no differences found between the SD of the Thermoscan and the PA and rectal temperatures.
- 

### **Terndrup and Milewski<sup>7</sup>** (quality of study – moderate)

#### MODEL

- Mean differences between rectal and uncorrected First Temp and Thermoscan were identical, mean  $1.2^{\circ}\text{C}$ , SD  $0.4^{\circ}\text{C}$ .
- However the corrected mean of the First Temp (mean  $37.6^{\circ}\text{C}$ , SE  $0.05^{\circ}\text{C}$ ) was found to be significantly lower than the mean rectal temperature (mean  $38.0^{\circ}\text{C}$ , SE  $0.05^{\circ}\text{C}$ ; p value not given). There was no significant difference found between the mean rectal temperature and the mean Thermoscan temperature (mean  $38.0^{\circ}\text{C}$ , SE  $0.05^{\circ}\text{C}$ ). It should be remembered that  $0.8^{\circ}\text{C}$  was added to First Temp and  $1.2^{\circ}\text{C}$  was added to Thermoscan.
- Sensitivities and specificities were provided for the 2 models at selected rectal temperatures. The sensitivity of Thermoscan was shown to be higher than First Temp: conversely the specificity of First Temp was higher than Thermoscan. ROC curve analysis found that the Thermoscan had a larger area under the curve than First Temp (actual figures or statistical testing not provided).
- Given that the article has not provided any statistical testing, we can not conclude that one model is more accurate than the other.

#### CHILD'S AGE

There was no significant difference in rectal minus uncorrected tympanic temperatures for the various age groups (using correlation coefficients).

---

### **Terndrup and Wong<sup>8</sup>**: (quality of study – moderate)

#### OTITIS MEDIA

- Acute otitis media was seen in 35.5% (176/496) of ears.
- The mean temperature of ears with otitis media is statistically higher than ears without otitis media both rectally and with ITT (rectal temperature mean  $38.5^{\circ}\text{C}$ , SD  $1.0^{\circ}\text{C}$ , vs  $38.0^{\circ}\text{C}$ , SD  $0.9^{\circ}\text{C}$   $p < 0.05$ ; ITT temperature mean  $38.2^{\circ}\text{C}$ , SD  $1.1^{\circ}\text{C}$  vs  $37.6^{\circ}\text{C}$ , SD  $0.9^{\circ}\text{C}$ ,  $p < 0.05$ ).
- However, there was no significant difference in the bias (RT – ITT) of ears with otitis media and in ears without otitis media.
- Elevated temperatures recorded in ears with otitis media are likely to be due to systemic temperature increases rather than in local increases.
- Children with unilateral otitis media, there was no significant difference between the ITT of the infected vs uninfected ear.

#### CHILD'S AGE

The article reports that age less than 29 months had little influence in the relationship between ITT and RT – ie the correlation was not significantly different for age groups below 29 months. There was a significant reduction in the correlation coefficient for subjects older than 29 months compared with the groups younger than 29 months. The mean biases of different age groups are similar  $0.3^{\circ}\text{C}$  –  $0.4^{\circ}\text{C}$ ; but there is more variability in the age groups above 29 months ( $0.7^{\circ}\text{C}$  to  $0.8^{\circ}\text{C}$  vs  $0.4^{\circ}\text{C}$  to  $0.6^{\circ}\text{C}$ ), which may account for the difference in correlation coefficients. Given that the numbers in the groups are small ( $n < 50$  in each group), it is difficult to draw statistically or clinically significant conclusions from these findings.

---

### **Weiss and Richards<sup>9</sup>** (quality of study – low)

#### MODE OF THERMOMETER

- There does not appear to be any clinically significant difference between the mean bias (axillary temperature minus rectal temperature) of the axillary thermometer used in predictive mode and the axillary thermometer used in monitor mode.