



Centre for Clinical Effectiveness

Evidence Request 2007-005a

Infrared tympanic thermometers versus digital/electronic or infrared axillary thermometers for temperature recording 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 effect the accuracy of these two types of thermometers. Please refer to Evidence Request 2007-005b for this question.

Clinical Question: In children and infants does infrared tympanic thermometer or electronic/digital or infrared axillary thermometer provide a more accurate temperature measurement?

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: There were a total of 106 references found and six cohort studies were reviewed. The studies were of variable quality with some methodological problems, but had consistent results.

Of the five studies included which used electronic/digital axillary thermometers, four studies found infrared tympanic thermometers to be more accurate than axillary thermometers. Of these studies, three found the differences to be statistically significant, and one did not provide statistical testing but the differences were clinically important. One study found no differences between the two types of thermometers⁴. It should be noted that this study was conducted on neonates

The one study which used infrared axillary thermometers found the infrared tympanic thermometer to be statistically better able to detect fever than infrared axillary thermometer.

Conclusions: The evidence suggests that in children, one month to 15 years, infrared tympanic thermometers are more accurate at temperature measurement than axillary thermometers. There is no evidence that in neonates there is any difference between the two types of thermometer.

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 recording children's temperatures.

She also requested information about the factors that effect the accuracy of these two types of thermometers. Please refer to Evidence Request 2007-005b for this part of her question.

Clinical Question

In children and infants does infrared tympanic thermometer or electronic/digital or infrared axillary thermometer provide a more accurate temperature measurement?

Methods

Study Selection Criteria

Inclusion and exclusion criteria

Patient	Inclusion: <ul style="list-style-type: none">Children and infants (0-15 years) Exclusion: <ul style="list-style-type: none">Preterm babies (<37 weeks gestational age)Infants in heated bassinets, cribs, incubators or radiant warmersA population of solely hypothermic or hyperthermic participantsChildren being cooled, warmed or rewarmedCranio-facial or ear canal malformations
Intervention	Inclusion: <ul style="list-style-type: none">Infrared tympanic thermometer Exclusion: <ul style="list-style-type: none">All other (non-infrared) tympanic thermometers
Comparison	Inclusion: <ul style="list-style-type: none">Axillary thermometer (digital/electronic)Axillary thermometer (infrared) Exclusion: <ul style="list-style-type: none">All other types of axillary thermometersThermometers which use a skin temperature sensor
Outcomes	Inclusion: <ul style="list-style-type: none">Temperature Exclusion: <ul style="list-style-type: none">All other outcomes
Reference standard	Core body temperature (arterial blood, bladder or rectal) using digital/electronic or glass-mercury thermometer
Study design	Inclusion: <ul style="list-style-type: none">Prospective cohort studies where tympanic, axillary and core body temperature have been measured in all participants.Systematic review of the above study type Exclusion: <ul style="list-style-type: none">All other studies
Publication details	Inclusion: Studies in English Exclusion: Studies in languages other than English

Search Strategy

Evidence Source	Date of Search
<u>Databases:</u>	
<ul style="list-style-type: none">All EBM Reviews (Ovid)*	27 th Feb 2007
<ul style="list-style-type: none">Medline (Ovid) (from 1966)	27 th Feb 2007
<ul style="list-style-type: none">CINAHL (Ovid) (from 1982)	27 th Feb 2007
<ul style="list-style-type: none">The Cochrane Library**	27 th Feb 2007
<u>Relevant websites</u>	

• The Joanna Briggs Institute: www.joannabriggs.edu.au	21 st Nov 2006
• US National Guideline Clearinghouse: www.guidelines.gov	21 st Nov 2006
• UK National Institute for Clinical Excellence: www.nice.org.uk	21 st Nov 2006
• Scottish Intercollegiate Guidelines: www.sign.ac.uk	21 st Nov 2006
• New Zealand Guidelines Group: www.nzgg.org.au	21 st Nov 2006
• NH&MRC guidelines www.nhmrc.gov.au	21 st Nov 2006
• Guidelines International Network: www.g-i-n.net	21 st Nov 2006
• Trip Database: www.tripdatabase.com	21 st Nov 2006
• Google: www.google.com.au	21 st 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 HTAs

Search Terms

Patient	exp Child/ OR exp Pediatrics/ OR exp Infant/ OR (child\$ or pediatric\$ or paediatric\$ or infant\$).mp.
Intervention	((thermomet\$ or temperature\$).mp. OR exp Body temperature/) AND (tympanic\$ or ear\$ or aural).mp.
Comparison	axilla\$.mp.
Outcomes	-

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 membrane⁴.
- **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 familiar¹³.

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 temperature¹⁴.

- **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 – specificity¹¹. A global assessment of the performance of the test (sometimes called diagnostic accuracy) is given by the area under the receiver operating characteristic curve¹¹ (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 test¹⁰. 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 test¹⁰. 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

There were a total of 106 references found with the above search terms. After reviewing the title, abstract or full text, one systematic review⁷ and eight cohort studies^{1-6,8,9} were found which addressed the research question. We were unable to obtain the systematic review (despite contacting the author several times)⁷. One cohort study was excluded because it failed to provide adequate information on methodology or technique used⁹. Another study was excluded as we were unable to interpret the results and communication with the authors did not provide clarification⁸.

Five studies compared electronic/digital axillary thermometers to infrared tympanic thermometers^{1,3-6}. Only one study compared an infrared axillary thermometers to an infrared tympanic thermometers².

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

Electronic/digital axillary thermometers and infrared tympanic thermometers

Nimah et al¹ conducted a study between 2000 and 2002. They enrolled 37 children under seven years of age (mean age 20.0 months) who were in an intensive care unit and required an indwelling bladder catheter. The exclusion criteria were long (see Appendix 1) but of note was that children with otitis media were excluded.

Infrared tympanic thermometer (ITT) temperatures were taken with the Braun Thermoscan, IRT 3020 and IRT 3520 using the ear tug. A Turbo-Temp digital thermometer was used for axillary temperature. No information was provided about the modes used. Bladder temperature was the reference standard and was monitored using an indwelling RSP Foley catheter with 400 series thermistor. In addition, rectal temperature was measured with an indwelling Mon-a-therm rectal probe. Measurements were taken every hour, except if there was an increase of 1°F (0.56°C) and then they were taken every five minutes until the core temperature was constant (<0.5°F (0.28°C) of variability for three consecutive measurements taken five minutes apart).

This study was of moderate quality. One potential area for concern was that although 387 patients were eligible for enrolment, only 61 were approached for consent. It was not documented why so few were approached, but this may be because of the long list of exclusion criteria. All temperature measurements were taken by the same person and therefore the measurements could not be independent or blind.

The study found the mean difference (bladder temperature minus test measure) of the ITT was 0.03°F, standard deviation 0.72°F (0.0167°C, SD 0.40°C). The mean difference between bladder and axillary thermometer was 1.25°F, standard deviation 0.87°F (0.694°C, SD 0.48°C). The bladder temperature was higher than ITT or axillary. Although these differences do appear clinically significant, the statistical significance was not reported.

ITT measurements had better agreement with core measurements during increasing and decreasing temperature cycles. The temperature measurements were separated into 'steady state measurements', 'increasing 5-minute measurements', 'plateau measurements' and 'decreasing five-minute measurements'. The mean difference of the ITT measurements was smaller than the mean difference of the axillary measurements. When temperature was changing rapidly, as in the five-minute measurement periods, there were smaller differences seen in the ITT. For example in the 'decreasing five-minute measurements' ITT mean difference 0.16°F, standard deviation 0.56°F (0.09°C, SD 0.314°C), compared to axillary mean difference -1.95°F, standard deviation 0.9°F (-1.09°C, SD 0.5°C). It was not documented if these differences were statistically significant.

Sensitivity and specificity were 80% and 81% for ITT and 40% and 98% for axillary temperature (at core temperature 100.4°F (38°C)). Receiver operator characteristic (ROC) curve analysis revealed that ITT measurements performed well, with an area under the curve (AUC) of 0.906 (95%CI 0.880, 0.924) in comparison with axillary temperature 0.847 (95%CI 0.821, 0.873). In summary ITT was found to be statistically significantly more accurate than axillary temperature measurement.

Kongpanichkul & Bunjongpak³ conducted a study in 1998 of 200 infants 0–48 months, with birth weight at least 3 kgs, attending the "Paediatric Department" of a hospital in Thailand.

ITT temperature was taken using the Welch Allyn model 9000 on rectal equivalence setting with the use of the ear tug. Axillary temperatures were taken using the Terumo Corporation model C202. Mode used was not documented. The reference measurement was taken using a calibrated glass-mercury rectal thermometer. A liquid crystal forehead thermometer was also used and compared to rectal temperature, but this part of the study was not relevant for our review. Measurements were obtained by each device three times and the means analysed. Two examiners performed the measurements on each child. There were four examiners in total and they were all instructed to use the types of devices accurately.

The study was of moderate quality. It was not documented how the sample was chosen. There were two examiners taking the four measurements, so they would each know at least one of the other measurements, therefore they could not be independent or blind.

This study reported that for fever (rectal temperature $\geq 38^\circ\text{C}$) the ITT had a sensitivity of 71.9% and a specificity of 98.1%; compared with axillary thermometer which had a sensitivity of 67.7% and a specificity of 99.0%. For high fever (rectal temperature $\geq 39^\circ\text{C}$) ITT had a sensitivity of 77.4% and a specificity of 99.4%; compared with axillary temperature

which had a sensitivity of 67.7% and a specificity of 98.8%. At high fever, the ITT does appear to have a clinically important higher sensitivity than the axillary thermometer (77.4% vs 67.7%). However, as the AUC and CIs have not been reported, it is unclear if there is any statistically significant difference between the two methods of measurement.

Weiss et al⁴ conducted a study which was published in 1994 (not detailed when the study was undertaken) and included 34 neonates, 38–42 weeks gestational age, weighing >2500g, in a newborn nursery. The neonates were aged between two hours and four days at the time of data collection.

Tympanic temperatures were taken using the Thermoscan professional model and the Thermoscan consumer model. Three left ear and three right ear measurements were taken with each model. Both models were used in actual predictive mode and the ear tug was used. Axillary temperature was taken using an IVAC 2080 electronic thermometer, which was set in predictive mode. The reference in this study was an IVAC 2080 electronic rectal thermometer, which was set in predictive mode. All instruments were checked for calibration at the start and completion of data collection.

This study was of low quality. Only limited information has been provided about the study participants, that is age and weight. It has not been documented how the sample was chosen. Sixty sets of temperature measurements were collected; up to three sets of recordings taken at least two hours apart were permitted for each neonate. The taking of 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. Fourteen temperature measurements were taken which would have taken some time. Therefore there would be a time delay between the first measurement taken (tympanic) and the last measurement (rectal). The potential problem with this is that the temperature may have changed in this time period.

The mean difference (rectal temperature minus test measure) of the Thermoscan, left ear was 0.4°C, standard deviation 0.4°C. The mean difference for Thermoscan, right ear was also 0.4°C (SD 0.3°C), consumer model in both ears was 0.6°C (SD 0.4°C). The mean difference between rectal and axillary thermometer was 0.4°C (SD 0.2°C). The rectal temperature was higher than ITT or axillary. When machines were checked for calibration at completion of data collection, the professional model tympanic thermometer and the axillary thermometer were found to be accurately calibrated. The consumer model read 0.2°C lower at the completion. Temperature differences reported relating to the consumer model should, therefore, be interpreted with consideration for this discrepancy. In summary, no significant differences were found between the accuracy of the ITT measurements and the axillary temperature measurements.

Romano et al⁵ conducted a study between 1991 and 1992, included 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 the tympanic membrane, presence of tympanostomy tubes or contraindication to the use of indwelling rectal tubes.

Tympanic temperatures were taken using 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. Digital axillary temperatures were taken using an IVAC Temp Plus II thermometer. The mode used was not specified. The reference measurement was taken using an Opticath PA catheter inserted using standard procedures. The study also used a Mon-A-Therm rectal probe.

This study was of moderate quality. Potential areas of concern include that tympanic and PA temperatures were measured in degrees Celsius; but axillary temperatures were measured in degrees Fahrenheit and converted to Celsius. There is a potential for error in the conversion. Tympanic thermometers were calibrated but there was no mention of calibration of the axillary thermometer. The authors had training in using the tympanic machines, but not with the axillary thermometer. These issues could have led to increased errors with the axillary temperature measurements. Limited information was provided about the participants. All temperature measurements were taken by the same person, therefore could not be independent or blind.

The mean difference (PA temperature minus test measure) of the First Temp ITT was -0.06°C (SD 0.58°C) and for the Thermoscan ITT this was -0.13°C (SD 0.39°C). That is to say tympanic temperature was higher. The mean difference between PA and axillary thermometer was 0.69°C (SD 0.60°C). In this instance the PA temperature was higher. Statistical testing found that the tympanic temperatures were not significantly different to the PA temperatures, but that the axillary temperatures were (paired t-test, $p < 0.0001$). In summary the study found that ITT measurements were statistically significantly more accurate than the axillary temperature measurements.

Muma et al⁶ conducted a study between 1986 and 1987 of 224 children less than 3 years old (mean age 12.4 months) who presented to the emergency department. Children who were immunocompromised, were receiving chemotherapy, or had rectal trauma, infection or anomalies were excluded.

ITT temperatures were taken with the FirstTEMP thermometer set in predictive rectal mode. It has not been documented if an ear tug was used. Axillary temperature was taken with the Diatek 500. Mode used was not specified. The reference used in this study was the Diatek 500 used rectally. Mode used was not specified.

This study was of moderate quality. Limited details of included participants were given. Children were sequentially enrolled into the study, how they were chosen was not stated. Information was not provided on how the measurements were taken or who took them. The probe was inserted only a few millimetres due to its relative large diameter of 8 mm compared with the 3 mm paediatric probe; it is not clear why the smaller paediatric probe was not used. This larger probe may have reduced the accuracy of the tympanic measurements.

This study reported that the mean difference of the ITT (rectal temperature minus ITT) was 0.71°C (SD 0.62°C); the difference of the axillary thermometer (rectal temperature minus axillary temperature) was 1.52°C (SD 0.67°C). The rectal temperature was higher than the ITT or axillary temperature measurements. We undertook statistical testing (a paired sample t test, using STATA 8, www.stata.com) on these figures and the difference between the two methods was found to be statistically significant ($p < 0.001$). Fever was defined as rectal temperature $\geq 38^\circ\text{C}$, tympanic temperature $\geq 38^\circ\text{C}$ and axillary temperature $> 37.3^\circ\text{C}$. Sensitivity and specificity were 55% and 100% for ITT and 48% and 95.4% for axillary. The ITT was better able to detect fever and non-fever than the axillary thermometer. Combining this result with the mean differences reported above, we found that ITT measurement more accurate than axillary temperature measurement.

Infrared axillary thermometers and infrared tympanic thermometers

A study by Jean-Mary et al² 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) attending 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. Axillary temperature was measured using an infrared thermometer, type not specified. Mode used was not documented for either the ITT or the axillary thermometers, but the article stated that both "thermometers reflect 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" to each value. The reference standard was an Ivac digital rectal thermometer. 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.

This study reported that the mean difference of the ITT (ITT minus rectal temperature) was -0.24°F (-0.133°C). The axillary temperature mean difference (axillary minus rectal temperature) was -0.33°F (-0.183°C). Standard deviations were not provided. The rectal temperature was higher than ITT or axillary temperature. These differences "were not statistically significantly different from each other with regards to means or standard deviations". As 1°F (0.56°C) was added to the tympanic and axillary temperatures, the uncorrected tympanic and axillary differences were -0.693°C and -0.743°C respectively. This study reported that sensitivity and specificity of the tympanic thermometer to detect fever (rectal temperature $\geq 100.4^\circ\text{F}$ (38.0°C)) were 68.3% and 94.8% respectively; for the axillary thermometer these were 63.5% and 92.6% respectively. ROC curve analysis revealed that ITT performed statistically significantly better than axillary thermometers; with an AUC 0.88 compared to 0.81 ($p < 0.04$, CIs not provided). In summary, the ITT thermometer was shown to be statistically better able to detect fever than infrared axillary temperature thermometer.

Discussion

The studies included in this review were of variable quality with all having some methodological issues. Two studies did not provide specific inclusion and exclusion criteria^{2,4}, and some studies provided limited descriptors of the study participants^{2,4,5,6}. The problem with these issues is that we are unable to identify 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 (and in one study two assessors³) recorded all the measurements. Although this is a flaw of study design it is difficult to know how this would affect the results.

The studies included used a variety of different types of ITT and axillary thermometers. It is difficult to say if the accuracy of these devices varies. Evidence Request 2007-004b evaluates this. The emergency department of Casey Hospital uses Genius tympanic thermometers, which was tested in one of the studies⁵. For the axillary thermometer they use Welch Allyn, Sure Temp and Sure Temp plus. These thermometers were not assessed in any of the included studies. The findings of this report need to be viewed taking into consideration that different models of machine may have different accuracy levels and we are only able to report on devices included in the studies found.

Of the five studies included which used electronic/digital axillary thermometers, four studies found ITT to be more accurate than axillary thermometers. Of these studies, three found the differences to be statistically significant^{1,5,6}, and one did not provide statistical testing but the differences were clinically important³. One study found no differences between the two types of thermometers⁴. It should be noted that this study was conducted on neonates who have a small ear canal which the authors suggest may have been a source of measurement error⁴. The tympanic tip used was 7.4 mm compared to the ear canal size of the neonate which is approximately 4 mm.

The one study which used infrared axillary thermometers found the ITT thermometer was statistically better able to detect fever than infrared axillary temperature thermometer.

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 times¹². Therefore this site of core temperature is not used as frequently¹².

Conclusions

The evidence suggests that in children, one month to 15 years, infrared tympanic thermometers are more accurate at temperature measurement than axillary thermometers. There is no evidence that in neonates there is any difference between the two types of thermometer.

References

1. Nimah, M.M., et al., *Infrared tympanic thermometry in comparison with other temperature measurement techniques in febrile children*. Pediatric Critical Care Medicine, 2006. **7**(1): p. 48-55.
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. Kongpanichkul, A. and S. Bunjongpak, *A comparative study on accuracy of liquid crystal forehead, digital electronic axillary, infrared tympanic with glass-mercury rectal thermometer in infants and young children*. Journal of the Medical Association of Thailand, 2000. **83**(9): p. 1068-76.
4. Weiss, M.E., D. Poeltler, and I. Gocka, *Infrared tympanic thermometry for neonatal temperature assessment*. JOGNN - Journal of Obstetric, Gynecologic, & Neonatal Nursing, 1994. **23**(9): p. 798-804.
5. Romano, M.J., et al., *Infrared tympanic thermometry in the pediatric intensive care unit*. Critical Care Medicine, 1993. **21**(8): p. 1181-5.
6. 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.
7. Duce, S., *A systematic review of the literature to determine optimal methods of temperature measurement in neonates, infants and children (Structured abstract)*. Database of Abstracts of Reviews of Effects, 1996(4): p. 1-124.
8. 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.
9. El-Radhi, A.S. and S. Patel, *An evaluation of tympanic thermometry in a paediatric emergency department*. Emergency Medicine Journal, 2006. **23**(1): p. 40-1.
10. Altman, D.G. and J.M. Bland, *Diagnostic tests. 1: Sensitivity and specificity*. BMJ, 1994. **308**(6943): p. 1552.
11. Altman, D.G. and J.M. Bland, *Diagnostic tests 3: receiver operating characteristic plots*. BMJ, 1994. **309**(6948): p. 188.
12. 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.
13. Craig, J.V., et al., *Infrared ear thermometry compared with rectal thermometry in children: a systematic review..* Lancet, 2002. **360**(9333): p. 603-9.
14. 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.

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

Appraisal of included studies

Study: Nimah et al¹, 2006

Description of study

Patients	Critically ill children <7 years of age with an indwelling bladder catheter
N	37
When conducted	Study conducted between Oct 2000 and Oct 2002
Setting	Paediatric or cardiac intensive care unit in an US hospital
Infrared tympanic	
1. Type	1. The Braun Thermoscan, IRT 3020 and IRT 3520
2. Mode used	2. Mode not specified
3. Ear tug	3. Ear tug was used
4. Otitis media	4. Patients with otitis media were excluded
5. Calibration	5. Calibration using water bath methodology was undertaken by the manufacturer before use in the study
6. Technique	6. The probe was placed snugly into the external auditory canal
Axillary	
1. Type	1. A Turbo-Temp digital thermometer
2. Mode used	2. Mode not specified
3. Calibration	3. Calibration not specified
4. Technique	4. Manufacturer's instructions were followed
Outcomes	<p>Temperatures recorded during febrile and nonfebrile periods. Measurements taken every hour except if there was an increase of 1°F (0.56°C) and then they were taken every 5 minutes, until the core temperature was stable. Measurements were taken from all sites in a rapid sequential manner. Research nurses received extensive training on all temperature monitoring devices.</p> <p>Fever defined as $\geq 100.4^{\circ}\text{F}$ (38.0°C)</p> <p>Sensitivity, specificity, positive and negative predictive values were calculated for temperature cutoffs between 99.0°F (37.2°C) and 101.5°F (38.6°C)</p>
Reference standard	
1. Type	1. Bladder temperature measured with an indwelling RSP Foley catheter with 400 series thermistor;
2. Mode used	2. Mode not applicable
3. Calibration	3. Standard calibration was performed by manufacturer

4. Technique	4. The bladder temperature monitor was set every morning to alarm if the patient's temperature increased above 1°F (0.56°C) above baseline Rectal temperature measured with an indwelling Mon-a-therm rectal probe
Inclusion Criteria	As above
Exclusion criteria	Required induced hypothermia; required overhead warmer during the study; were diagnosed with diabetes insipidus; had urine output of <1 mL/kg/hr; were diagnosed with otitis media or had drainage from either ear or mastoiditis; were wearing a hearing aid in the ear in which the temperatures were to be taken; had any contraindication for rectal probe; were diagnosed with any anatomic abnormalities that would affect temperature measurements of the ear, forehead, axilla or rectum; had a platelet count of <50,000/microL or absolute neutrophil count of <500/microL.

Study: Jean-Mary et al², 2002

Description of study

Patients	Children 3-36 months old
N	198
When conducted	Not detailed when the study was conducted
Setting	An outpatient department for either 'well child' or acute illness visits in the US
Infrared tympanic	
1. Type	1. Type not specified.
2. Mode used	2. Mode used not documented, but the manufacture's handbook states that the temperature reflects an approximation of oral temperature.
3. Ear tug	3. Ear tug not specified
4. Otitis media	4. Otitis media not specified
5. Calibration	5. Calibrated according to hospital protocol or manual
6. Technique	6. Used as described in the user manual. 1°F (0.556°C) added to bring in line with rectal temperature.
Axillary	
1. Type	1. Infrared thermometer, type not specified
2. Mode used	2. Mode used not documented, but the manufacture's handbook states that the temperature reflects an approximation of oral temperature.
3. Calibration	3. Calibrated according to hospital protocol or manual
4. Technique	4. Used as described in the user manual 1°F (0.556°C) added to bring in line with rectal temperature.

Outcomes	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°C)
Reference standard	
1. Type	1. An Ivac digital rectal thermometer
2. Mode used	2. Mode used not specified
3. Calibration	3. Calibrated according to hospital protocol or manual
4. Technique	4. Technique not specified
Inclusion Criteria	Nil other than specified above
Exclusion criteria	Contraindication to rectal temperature measurement or known hypothalamic dysfunction

Study: Kongpanichkul & Bunjongpak³, 2000

Description of study

Patients	Children aged 0–48 months with birth weight at least 3 kgs.
N	200
When conducted	Study conducted between Aug and Oct 98
Setting	The “Paediatric Department” of a hospital in Thailand
Infrared tympanic	
1. Type	1. Welch Allyn model 9000
2. Mode used	2. Rectal equivalence mode
3. Ear tug	3. Ear tug used
4. Otitis media	4. Otitis media not specified
5. Calibration	5. Calibration not specified
6. Technique	6. The thermometer was placed in the right external auditory canal with a tight seal
Axillary	
1. Type	1. Terumo Corporation model C202.
2. Mode used	2. Mode not specified
3. Calibration	3. Calibration not specified
4. Technique	4. The probe was placed next to the skin under the axilla

Outcomes	Measurements were obtained by each device three times and the means analysed. Two examiners performed the four measurements for each child. There were four examiners in total and they were all instructed to use the four types of devices accurately Fever defined as $\geq 38.0^{\circ}\text{C}$; high fever defined as $\geq 39.0^{\circ}\text{C}$
Reference standard	
1. Type	1. Glass–mercury rectal thermometer
2. Mode used	2. Mode not applicable
3. Calibration	3. The article states that the glass-mercury thermometer was calibrated, but does not give specific details
4. Technique	4. Placed into the rectum at a depth of 2 cm in neonates and 3 cm in older infants and children — left in place for one minute
	A liquid crystal forehead thermometer was also used.
Inclusion Criteria	Children aged 0–48 months with birth weight at least 3 kgs
Exclusion criteria	Infants and children with abnormal otic, rectal structure, diarrhoea and contagious skin disease

Study: Weiss et al⁴, 1994

Description of study

Patients	38–42 weeks gestational age, weighing >2500gms. Aged between 2 hours and 4 days at the time of data collection.
N	34
When conducted	Not clear when the study was conducted
Setting	Newborn nursery of a tertiary level perinatal centre in the US
Infrared tympanic	
1. Type	1. Thermoscan professional model Instant and the Thermoscan consumer model Instant
2. Mode used	2. Used in actual predictive mode.
3. Ear tug	3. Ear tug used.
4. Otitis media	4. Otitis media not specified
5. Calibration	5. All instruments were checked for calibration at the start and completion of data collection.
6. Technique	6. Measurements were taken by a single trained data collector.
Axillary	
1. Type	1. An IVAC 2080 electronic thermometer
2. Mode used	2. Used in predictive mode

3. Calibration	3. All instruments were checked for calibration at the start and completion of data collection
4. Technique	4. A left axillary temperature was taken
Outcomes	Sixty sets of temperature measurements were collected; up to three sets of recordings taken at least 2 hours apart were permitted for each neonate. Thermoscan professional model Instant; 3 left ear and 3 right ear measurements were taken; then 3 left ear and 3 right ear measurements using the Thermoscan consumer model Instant Infrared thermometer
Reference standard	
1. Type	1. An IVAC 2080 electronic rectal thermometer
2. Mode used	2. Used in predictive mode
3. Calibration	3. All instruments were checked for calibration at the start and completion of data collection
4. Technique	4. The probe was inserted to a depth of three fourths of an inch
Inclusion Criteria	Full-term neonates weighing at least 2500 grams
Exclusion criteria	Nil documented.

Study: Romano et al⁵, 1993

Description of study

Patients	6 months to 15 years requiring pulmonary artery catheter monitoring.
N	20
When conducted	Study conducted between July 1991 and March 1992
Setting	ICU of a tertiary paediatric hospital in the US
Infrared tympanic	
1. Type	1. Genius First Temp and Thermoscan Pro-1
2. Mode used	2. Used in core mode
3. Ear tug	3. An ear tug manoeuvre was used for the Thermoscan Pro-1 as recommended by the manufacturer.
4. Otitis media	4. Otitis media not specified
5. Calibration	5. Calibrated using a manufacturer-supplied "black-body" temperature reference source before patient enrolment and at the midpoint of the study.
6. Technique	6. Measurements were taken in degrees Celsius
Axillary	
1. Type	1. Temp Plus II, IVAC
2. Mode used	2. Mode used not specified

3. Calibration	3. Calibration not specified
4. Technique	4. Right axilla used. Measurements were taken in degrees Fahrenheit and converted to degrees Celsius
Outcomes	All measurements were taken 'simultaneously' by one of the authors, each of whom was instructed by a manufacturer's representative on the proper use (for the tympanic thermometers).
Reference standard	
1. Type	1. Opticath pulmonary artery catheters were used
2. Mode used	2. Mode not applicable
3. Calibration	3. Calibration not specified
4. Technique	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.
Inclusion Criteria	Admitted to paediatric ICU and requiring pulmonary artery catheter
Exclusion criteria	Inability to visualise tympanic membrane, presence of tympanostomy tubes or contraindication to the use of indwelling rectal tubes.

Study: Muma et al⁶, 1991

Description of study

Patients	Children less than 3 years old
N	224
When conducted	Study conducted between April 86 and April 87
Setting	Paediatric emergency department in US hospital
Infrared tympanic	
1. Type	1. The FirstTEMP thermometer
2. Mode used	2. Used in predicted rectal mode.
3. Ear tug	3. Ear tug not specified
4. Otitis media	4. The presence of otitis media was noted
5. Calibration	5. Compared biweekly with a glass-mercury thermometer.
6. Technique	6. The probe was only inserted a few millimetres due to its relative large diameter of 8mm compared with the 3 mm paediatric probe. Not reported if any training given
Axillary	

1. Type	1. Diatek 500 electronic thermistor probe
2. Mode used	2. Mode not specified
3. Calibration	3. Calibration according to factory setting was performed daily for the first two weeks and biweekly thereafter. Also compared biweekly with a glass-mercury thermometer.
4. Technique	4. No information provided
Outcomes	Temperatures were taken sequentially at the three sites; order not detailed The definition of fever was rectal >38°C, axillary > 37.3°C and tympanic ≥38°C.
Reference standard	
1. Type	1. Diatek 500 Electronic thermistor probe used rectally
2. Mode used	2. Mode not specified
3. Calibration	3. Calibration according to factory setting was performed daily for the first 2 weeks and biweekly thereafter. Also compared biweekly with a glass-mercury thermometer.
4. Technique	4. No information provided
Inclusion Criteria	Children less than three years old presenting to the emergency department
Exclusion criteria	Children who were immunocompromised, were receiving chemotherapy, or had rectal trauma, infection or anomalies were excluded.

Quality of included 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 reported	Comments
Nimah et al ¹ 2006	Yes	Yes	No	Yes	Yes	No	No	Yes	387 patients were eligible for enrolment, but only 61 were approached for consent. It is not documented why so few were approached. All temperature measurements were taken by the same person, therefore these could not be independent or blind.
Jean-Mary et al ² 2002	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 (contraindications to rectal temperature measurement or 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.
Kongpanichkul & Bunjongpak ³ 2000	Yes	Yes	Not stated	Yes	Yes	Probably not	Probably not	Yes	It has not been documented how the sample was chosen There were two examiners taking the four measurements, so they would each know at least one of the other measurements, therefore could not be independent or blind.
Weiss et al ⁴ 1994	No	Some	No	Yes	Yes	No	No	No	We have limited information about the included participants as the only study participant descriptors are age and weight. It has not been documented how the sample was chosen. Sixty sets of temperature measurements were collected; up to three sets of recordings taken at least 2 hours apart were permitted for each neonate. The taking of 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. Fourteen temperature measurements were taken which would have taken some time. Therefore there may be a time delay between the first measurement taken (tympenic) and the last one (rectal) The potential problem with this is that the temperature may have changed between the first and last measurements.
Romano et al ⁵ 1993	Yes	Some	Yes	Yes	Yes	No	No	No	Limited information provided about the participants. Tympanic and pulmonary artery temperatures were measured in degrees Celsius; but axillary temperatures were measured in degrees Fahrenheit and converted to Celsius. Tympanic thermometers were calibrated but no mention of calibration of the axillary thermometer. The authors, who took the tympanic measurements, had

									<p>training in using the tympanic machines, but not with the axillary thermometer. All these could have lead to increased errors with the axillary temperature measurements.</p> <p>All temperature measurements were taken by the same person, therefore these could not be independent or blind. It does appear these were objective.</p>
Muma et al ⁶ 1991	Yes	Some	Probably	Yes	Yes	No	No	Yes	<p>Limited details of participants given.</p> <p>Children were sequentially enrolled into the study, how they were chosen is not stated.</p> <p>Information not provided on how the measurements were taken or who took them.</p> <p>The probe was inserted only a few millimetres due to its relatively large diameter of 8 mm compared with the 3 mm paediatric probe – not clear why they didn't use the smaller paediatric one. This may increase the difference between the core temperature and the tympanic measurements.</p>

Results of included studies:

Nimah et al¹ 2006 (quality of study - moderate)

This study found that:

- Mean difference (bladder temperature minus test measure) of the ITT was 0.03°F, standard deviation 0.72°F (*0.0167°C, SD 0.40°C*). The mean difference between bladder and axillary thermometer was 1.25°F, standard deviation 0.87°F (*0.694°C, SD 0.48°C*). The bladder temperature was higher than ITT or axillary. Although these differences do appear clinically significant, the statistical significance was not reported.
- ITT measurements had better agreement with core measurements during increasing and decreasing temperature cycles. The temperature measurements were separated into 'steady state measurements', 'increasing 5-minute measurements', 'plateau measurements' and 'decreasing five-minute measurements'. The mean difference of the ITT measurements was smaller than the mean difference of the axillary measurements. When temperature was changing rapidly, as in the five-minute measurement periods, there were smaller differences seen in the ITT. For example in the 'decreasing five-minute measurements' ITT mean difference 0.16°F, standard deviation 0.56°F (*0.09°C, SD 0.314°C*), compared to axillary mean difference -1.95°F, standard deviation 0.9°F (*-1.09°C, SD 0.5°C*). It was not documented if these differences were statistically significant.
- Sensitivity and specificity were 80% and 81% for ITT and 40% and 98% for axillary temperature (at core temperature 100.4°F (38°C)).
- Receiver operator characteristic (ROC) curve analysis revealed that ITT measurements performed well, with an area under the curve (AUC) of 0.906 (95%CI 0.880, 0.924) in comparison with axillary temperature 0.847 (95%CI 0.821, 0.873).
- In summary ITT was found to be statistically significantly more accurate than axillary temperature measurement.

Jean-Mary et al² 2002 (quality of study – moderate)

This study found that:

- Sensitivity, specificity and diagnostic accuracy of the axillary thermometer for rectal fever (100.4°F (38.0°C)) were 63.5%, 92.6% and 83.3% respectively; those for the tympanic thermometer were 68.3%, 94.8% and 86.4% respectively. (Diagnostic accuracy is the number of true positives plus the number of true negatives, divided by the total number of participants.)
- Receiver operator characteristic curve analysis revealed that ITT measurements performed statistically significantly better than axillary temperature measurement; with an area under the curve of 0.88 compared to 0.81 ($p < 0.04$) (CIs not given).
- The average tympanic difference (ITT minus rectal temperature was -0.24°F (*-0.133°C*) and the average axillary temperature difference (axillary temperature minus rectal temperature) was -0.33°F (*-0.183°C*). These differences "were not statistically significantly different from each other with regards to means or standard deviations". As 1°F (*0.56°C*) was added to the tympanic and axillary temperature, the uncorrected tympanic and axillary differences were -0.693°C and -0.743°C and respectively. Standard deviations were not provided.
- When subgroup analysis was conducted by age (<1 year and 1-3 years) and temperature (febrile and afebrile) the mean differences of the ITT measurements were less than the mean difference of the axillary measurements. The greatest difference was seen in the axillary, febrile, 1-3 year old group: mean difference for the axillary was -1.20°F, standard deviation 0.89°F (*-0.667°C, SD 0.494°C*), compared mean difference -0.36°F, standard deviation 1.00°F (*-0.20°C, SD 0.556°C*) for the ITT; the article reports this difference as being statistically significant ($p < 0.002$). There were statistically significant differences in the afebrile, under 1 year old group; ITT mean difference -0.10°F, standard deviation 1.00°F (*-0.056°C, SD 0.556°C*) compared with the axillary group mean difference 0.29°F, standard deviation 0.81°F (*0.161°C, SD 0.450°C*). These differences however were not considered, by the authors, to be clinically significant.

Kongpanichkul & Bunjongpak³ 2000 (quality of study - moderate)

This study found that:

- For fever ($\geq 38^\circ\text{C}$) tympanic thermometer had sensitivity of 71.9%, specificity of 98.1%, and accuracy of 85.5%; compared with axillary temperature which had a sensitivity of 67.7%, specificity of 99.0%, and accuracy of 84.0%.
- For high fever ($\geq 39^\circ\text{C}$) tympanic thermometer had sensitivity of 77.4%, specificity of 99.4%, accuracy of 96.0% compared with axillary temperature which had a sensitivity of 67.7%, specificity of 98.8%, accuracy of 94.0%.
- Given that area under the curve and CIs have not been given, we have to conclude that it is unclear if there is any statistically significant difference between the two methods of measurement. At high fever $\geq 39.0^\circ\text{C}$ the tympanic thermometer does appear have a higher sensitivity than the axillary thermometer which is a clinically important difference (77.4% vs 67.7%).

Weiss et al⁴ 1994 (quality of study - low)

NB All instruments were checked for calibration at the start and completion of data collection. The professional model ear thermometer and the electronic thermometer were found to be accurately calibrated. The consumer model met the calibration standard at the beginning of the study but read 0.2°C lower at its completion. Given that the professional model is the model intended for use by health professionals, the results below are for this model. Also in addition, the article states, "temperature differences reported relating to the consumer model should, therefore, be interpreted with consideration for this discrepancy".

This study found that::

- The mean difference (rectal temperature minus test measure) of the Thermoscan, left ear was 0.4°C , standard deviation 0.4°C . The mean difference for Thermoscan, right ear was also 0.4°C (SD 0.3°C), consumer model in both ears was 0.6°C (SD 0.4°C). The mean difference between rectal and axillary thermometer was 0.4°C (SD 0.2°C). The rectal temperature was higher than ITT or axillary.
- No significant difference was found between the accuracy of the ITT measurements and the axillary measurements. There was less variability between axillary temperatures than tympanic temperature measurement. There appears to be limited clinical importance of this. In addition, as the authors have not reported any statistical testing, the statistical significance is unclear.

Romano et al⁵ 1993 (quality of study – moderate)

This study found that:

- The mean difference (PA temperature minus test measure) of the First Temp ITT was -0.06°C (SD 0.58°C) and for the Thermoscan ITT this was -0.13°C (SD 0.39°C). That is to say tympanic temperature was higher. The mean difference between PA and axillary thermometer was 0.69°C (SD 0.60°C). In this instance the PA temperature was higher. Statistical testing found that the tympanic temperatures were not significantly different to the PA temperatures, but that the axillary temperatures were (paired t-test, $p < 0.0001$). In summary the study found that ITT measurements were statistically significantly more accurate than the axillary temperature measurements.
- The variability (SD) of both the axillary thermometer and the First Temp tympanic thermometer showed significantly greater variability than the SD of PA minus rectal temperatures ($p < 0.01$); whereas the Thermoscan model did not.

- The accuracy of the tympanic thermometers was shown to be better than the axillary thermometer. The variability of the axillary thermometer was poor, statistically greater than the variability of the PA and the rectal temperature. However the precision (variability) of the tympanic thermometers seemed to depend on model used, the Thermoscan had good precision, whereas the First Temp did not.

Muma et al⁶ 1991 (quality of study – moderate)

- Mean difference between rectal temperature and ITT (rectal temperature minus ITT) was 0.71°C (SD 0.62°C). rectal temperature and axillary temperature (rectal temperature minus axillary temperature) was 1.52°C (SD 0.67°C). (The axillary and tympanic temperature measurements were lower than the rectal temperature.) We undertook statistical testing (a paired samples t-test) on these figures and the difference between the two methods was found to be statistically significance ($p < 0.001$).
- Sensitivity and specificity for ITT (temp $\geq 38^\circ\text{C}$) were 55% and 100% respectively. These figures were 48% and 95.4% for axillary (axillary temp $> 37.3^\circ\text{C}$, rectal temp $\geq 38^\circ\text{C}$), CI not reported.
- The above results would indicate that ITT recording is more accurate than axillary temperature recording.