International Journal of Nursing
December 2014, Vol. 1, No. 2, pp. 115-134
ISSN 2373-7662 (Print) 2373-7670 (Online)

Published by American Research Institute for Policy Development
DOI: 10.15640/ijn.v1n2a9
URL: http://dx.doi.org/DOI: 10.15640/ijn.v1n2a9

Systematic Review and Meta-Analysis of the Diagnostic Accuracy of an Infrared Tympanic Thermometer for Use with Adults

Seong-Hi Park[1], Hee Seon Lee[2], Mi Jeong Kwack[3], Yun Su Cho[4] & Chang-Bum Kang[5]

Abstract

The aim of this study was to investigate the accuracy of infrared tympanic temperaturemeasurements compared to other methods of measuring temperature to detect fever in adults. Studies published between 1966 and 2012 from periodicals indexed in Ovid Medline, Embase, CINAHL, Cochrane, KoreaMed, NDSL, KERIS and other databases were selected using the following keywords: "infrared thermometer." QUADAS-II was utilized to assess the internal validity of the diagnostic studies. Selected studies were analyzed through a meta-analysis using MetaDisc 1.4.The analysis included fifteen diagnostic studies with high methodological quality, involving 1,468 subjects in total. The results of the meta-analysis showed that the pooled sensitivity, specificity and area under the curve (AUC) of infrared tympanic thermometers in adults over 18 years were 0.59 (95% CI 0.55, 0.63), 0.91 (95% CI 0.90, 0.92), and 0.85, respectively. For oral temperature readings, the pooled sensitivity was 0.61(95% CI 0.53, 0.68), the pooled specificity was 0.93 (95% CI 0.90, 0.95), and the AUC was 0.74. The meta-analysis results of infrared tympanic temperature in this study wereinterpreted in comparison to other non-invasive temperature measurement systems. The results of this study found that the diagnostic accuracy of infrared tympanic temperature measurements was not less than that of oral temperature measurements.

Keywords: Thermometer, Sensitivity, Specificity, Meta-analysis, Adults

[1]PhD, Assistant Professor, School of Nursing, Pai Chai University.
[2]PhD, Head Nurse, Department of Nursing Service, Korea University Anam Hospital, 73, Inchon-ro, Seongbuk-gu, Seoul, Korea, 136-705. TEL: 82-2-920-5903, FAX: 82-2-920-5204, E-mail: hsunangel@naver.com
[3]MPH, MSN, Manager, Quality Improvement Team, Korea University Anam Hospital
[4]PhD, Director, Department of Nursing Service, Korea University Anam Hospital
[5]PhD, Team Manager, Research Development Team, Korea Health Promotion Foundation

1. Introduction

Body temperature is a sensitive and reliable indicator of physiological integrity, a patient's physical status, and the existence and progression of disease. This important area of clinical data is used to make diagnoses and to provide treatment and nursing (Giuliano et al., 2000; Sohng et al., 2009). Getting an accurate measurement of core temperature, which reflects actual body temperature, is essential (Jeong & Yoo, 1997). However, since the human body shows different temperature readings in different areas of the body, it is difficult to represent the true core temperature (Lee & Kim, 2007).

In practice, rectal temperature has been used for years to estimate core temperature. However, due to the lack of a body temperature control system in the rectum, when a patient is in shock and has decreased blood flow, rectal temperature does not reflect the core temperature (Lee & Kim, 2007). An axillary temperature reading ofnewborns in clinical settingsshows a good match between core temperature and rectal temperature, but the same level of accuracy has not been found with oral or rectal temperature readings in adults.Moreover, not only can glass mercury thermometerseasily break, but it takes ten minutes to measure an adult's temperature, adding a time-consuming taskto the workload of nurses(Kozier, Erb, Blais, & Wilkinson, 1997).

The ideal method of measuring a person's temperature should be accurate, fast, and reflective of core temperature while being noninvasive, non-traumatic, user-friendly, and hygienic to everyone(van Staaij, Rovers, Schilder, & Hoes, 2003). Recently, infrared thermometry, which correlates better with core temperature, has been developed to solve the problems of conventional temperature measurement. Since the tympanic membrane receives the same artery blood from the hypothalamic area (the temperature control center), it is considered an ideal place to examine body temperature (Childs, Harrison, & Hodkinson, 1999).

However, some studies have shown that an infrared tympanic temperature reading may not be appropriate because ambient temperature can influence the tympanic drum (Yun & Lim, 2005), and because tympanic temperature may be lower than rectal temperature in patients with a high fever or outer tympanic infection (Wells, King, Hedstrom, & Youngkins, 1995).

Nevertheless, despite these drawbacks, infrared tympanic temperature has become more popular in clinical practice. In reaction to this popularity, numerous studies havebeen conducted to solve these problems, but have so far been limited to research neonates and children who may be easily influenced by external ambient temperatures from factors such as infections (Park, Park, & Kang, 2013). Therefore, it is necessary to examine whether infrared tympanic temperature readings can replace rectal, axilla, and oral temperature measurements in adults. The purpose of this study is to examine the diagnostic accuracy of infrared tympanic temperature in adults through a systemic review and meta-analysis, and to providesuggestionsfor nursing interventions.

2. Methods

2.1 Study Design

This study is a systematic review and meta-analysis of researchthat examined the accuracy of infrared tympanic temperature measurements in adult patients older than 18 years of age.This research was conducted based on information from the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy (Macaskill, Gatsonis, Deeks, Harbord, & Takwoingi, 2010) and Preferred Reporting Items for Systematic Reviews and Meta-Analyses (Moher, Liberati, Tetzlaff, & Altman, 2009).

2.2 Setting and Sample

- Participants: adult patients above 15 years of age who visited outpatient clinics or were admitted to hospitals;
- Index test: infrared tympanic temperature measurement;
- Comparator tests: axillary, rectal, oral, and core temperature measurements;
- Outcomes: true positive (TP), false positive (FP), false negative (FN), or true negative (TN);
- Reference standards: axillary, rectal, oral, and core temperature measurements;
- Type of studies: only diagnostic accuracy studies were included;

2.3 Ethical Consideration

This study did not need the approval of the institutional review board because the analyzed data are publicly available.

[

2.4 Measurement and Data Collection

2.4.1 Data Sources and Study Selection

A review of the literature through online databases was conducted on 31 July, 2013.KoreaMed, the National Discovery for Science Leaders (NDSL), and the Korean Education Research Information Service (KERIS) were used as the main Korean research databases. Additionally, the websites of the Korean Society of Nursing Science and the Korean Academy of Adult Nursing, Fundamentals of Nursing, Biological Nursing Science, Elderly Nursing, Nursing Education, Public Health Nursing, and Family Medicine were searched to include all Korean academic journals that deal with relevant fields. Ovid-Medline and Embase, the Cochrane Library, and CINAHL Plus with Full Text were used as the main international search databases. The keywords were derived from participants and interventions which were components of key questions.A search filter for diagnostic accuracy studies,a strategy proposed by the Scottish Intercollegiate Guidelines Network, was used for this research.

After removing overlapping first-search references, the title and abstract of each reference was reviewed for selection in accordance with the inclusion and exclusion criteria.The original papers of the remaining references after the primary ones had been excluded were then searched and again selected by applying the inclusion and exclusion criteria.The reference selection process was independently performed by two authors, and discussion and the application of the third party intervention principle was to be held in the case of any discordance; however, no conflicting opinionsarose between theinvestigators.

2.4.2 Risk of Bias in Included Studies

The methodological quality of selected studies was assessed using the Quality Assessment of Diagnostic Accuracy Studies-II (QUADAS-II) (Whiting et al., 2011). Two of the authors assessed bias independently; any disagreements or misunderstandings were resolved through discussion until a consensus was reached.

2.5 Dataanalysis

Relevant data such as the clinical characteristics of the study participants, the baseline, patient age, the selection criteria of subjects, the model of the thermometer, reference standard tests and febrile criteria, and the results of the infrared thermometer including TP, FP, FN, and TN were consolidated. Sensitivity,specificity, positive likelihood ratios, and the negative likelihood ratioswith 95% confidence intervals (CI) of each category were calculated using a 2-way Contingency Table Analysis.

Meta-analyses were performed using MetaDiSc, version 1.4 software. Using general principles for statistical modelling, the data were analyzed through the random effects model to identify heterogeneity including sensitivity, specificity, positive likelihood ratios, negative likelihood ratios, diagnostic odds ratios, and summary Receiver-Operating Characteristic (sROC) curves. Thestatistic value of an sROC curve is described as diagnostic test efficacy through the area under the curve (AUC) and the index Q^* value. AUC values may appear as follows: AUC=0.5: Non-informative; 0.5<AUC≤0.7: Less accurate; 0.7<AUC≤0.9: Moderately accurate; 0.9<AUC<1: Highly accurate; AUC=1: Perfect (Greiner, Pfeiffer, & Smith, 2000). In a Receiver Operating Characteristic (ROC) curve, the value for index Q^*,100%, corresponding toequal sensitivity and specificity may indicate 1(Walter, 2002). Confidence interval and effect size estimation through forest plots can be used to visually assess heterogeneity among studies.In this study,an I^2 test was also used to identify heterogeneity, while the chi-squared (χ^2) test was used to detect statistical heterogeneity. Here, values between 0% and 25% can be interpreted as unimportant heterogeneity, $25.0\%<I^2\leq75.0\%$ as moderate heterogeneity, and over 75% as considerable heterogeneity (Higgins & Thompson 2002).

3.Results

3.1 General Characteristics of Selected Studies

The initial literature research yielded 580 references, leaving 419 references once 161 duplicated studies had been excluded.

Selection criteria were studies that included an infrared thermometer as a diagnostic method with true positive, false positive, true negative, and false negative results in adult patients older than 18 years of age. The language of publication was not limited for studies reviewed.

The exclusion criteria were as follows:(a)studies performed using non-infrared thermometers, (b) evaluation studiesaimed at comparing diagnostic accuracy against the reference standard, (c) studies with subjects younger than 15 yearsof age, (d) studies in which the accuracy of diagnosis could not be estimated, (e) studies that did not have original authorship, (f) studies performed using thermometers in non-clinical settings, and (g) studies using animals or cadavers as subjects(Figure 1).

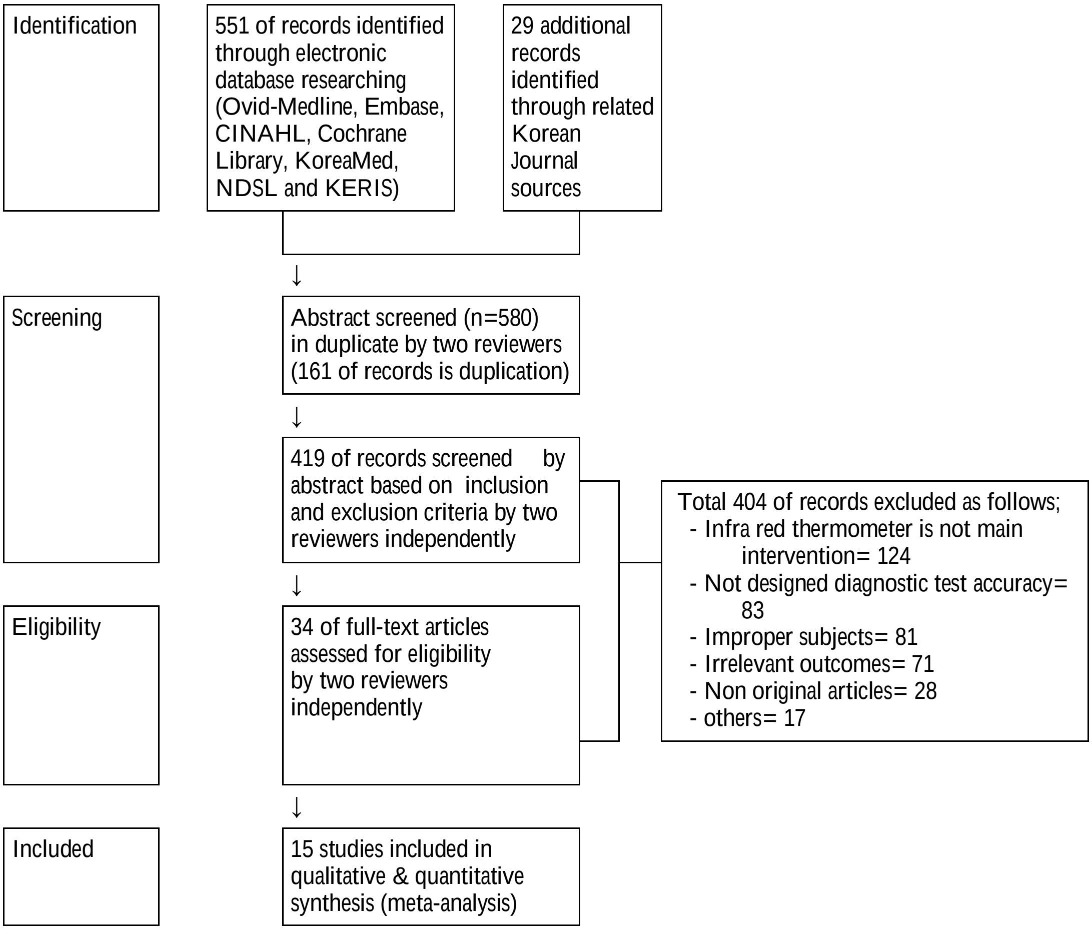


Figure 1: Flow Diagram of Article Selection

Fifteen studies involving 1,468 participants (2,179 total measurements) were selected to assess the diagnostic method for infrared tympanic thermometer measurement.

While the U.S. was the most active place of research on this topic with four studies place proceeding from that country, it was clear that this was a subject of worldwide interest as three of the studiescame out of Norway; two from Belgium; and one each from Australia, Canada, Denmark, Korea, Malaysia, and Sweden.

Studies were conducted for febrile patients in emergency rooms (Nordås, Leiren, & Hansen, 2005; Rajee & Sultana, 2006; Varney, Manthey, Culpepper, & Creedon, 2002; Yaron, Lowenstein, & Koziol-McLain, 1995), geriatric inpatient wards (Christensen, Christensen, & Matzen, 1998; Prentice & Moreland, 1999; Smitz, Giagoultsis, Dewe, & Albert, 2000; Smitz, Van de Winckel, & Smitz, 2009), cardiovascular inpatient wards, neurosurgery inpatient wards, and cancer inpatient wards. Patient ages ranged widely, with participants from18 years of age to elderly people. The number of subjects per study ranged from 21(Dzarr, Kamal, & Baba, 2009) to 213(Nordås et al., 2005), with seven relatively large-scale studies involving more than 100 subjects(Christensen et al., 1998; Duberg, Lundholm, & Holmberg, 2007; Nordås et al., 2005; Rajee & Sultana, 2006; Smitz et al., 2009; Valle, Kildahl-Andersen, & Steinvoll, 1999; Yaron et al., 1995). Moreover, 73.3% of the studies were large-scale ones. Even four relatively small-scale studies (Dzarr et al., 2009; Giuliano et al., 2000; Joo & Sohng, 2012; Petersen & Hauge, 1997) involving fewer than 100 subjects involved 200 to 300 distinct temperature measurements.The primary type of researchinvolved evaluation research to examinethe existence of fever using an infrared tympanic thermometer through a reference standard; only one case-control study compared febrile patients and controlled patients(Prentice & Moreland, 1999).

In all of the studies, temperature measurement was performed by trained nurses, with only Peterson& Hauge's (1997) research comparing the accuracy between trained and non-trained nurses.All studies were performed in compliance with the principles of temperature measurement. Ten of the studies included rectal temperature as the reference standard, while three used oral temperature (Prentice & Moreland, 1999; Rajee & Sultana, 2006; Skupski, Sonnenblick, Wagner, & Chervenak, 1995), and two employed core (pulmonary artery) temperature(Giuliano et al., 2000; Joo & Sohng, 2012).

Five studies compared infrared temperature measurement against other types of readings, two compared infrared and axillary temperature readings (Dzarr et al., 2009; Joo & Sohng, 2012), and four compared infrared and oral temperature readings (Dzarr et al., 2009; Giuliano et al., 2000; Nordås et al., 2005; Varney et al., 2002).

Some studies compared temperature readings according to clinical thermometer manufacturer (Giuliano et al., 2000; Nordås et al., 2005) and one investigated the accuracy of measurement between trained and non-trained nurses(Petersen & Hauge, 1997).While the definitions of “fever”differed slightly among studies, most papersindicated 38.0℃ as the cut-off point for fever (the exception Yaron et al., 1995, who used 38.5℃), and 37.0℃(Giuliano et al., 2000) or 37.6℃(Joo & Sohng, 2012) as the core temperature reading. Christensen et al. (1998) described the accuracy of diagnosis according to the cut-off point for fever (Table 1).

Table 1: Characteristics of Selected Studies

Year of publication	Authors	Location	Participants			Index test	Comparators	Reference standard	Fever criteria (°C)	2×2 Table				Value (95% Confidence interval)				
			Subjects	Age (yrs)	Total (N)					TP	FP	FN	TN	SN	SP	PLR	NLR	DOR
2012	Joo & Sohng	Korea	Cardiac care unit pts.	>18	83 (248)	ITT		PA	≥37.6	42	0	13	193	0.76 (0.70-0.76)	1.00 (0.98-1.00)	-	0.24 (0.24-0.31)	-
							A	PA	≥37.6	48	0	7	193	0.87 (0.81-0.87)	1.00 (0.98-1.00)	-	0.13 (0.13-0.20)	-
2009	Dzarr et al.	Malaysia	Cancer pts.	15-63	21 (300)	ITT		R	≥38.0	42	6	24	228	0.64 (0.55-0.69)	0.97 (0.95-0.99)	24.8 (11.23-61.76)	0.37 (0.32-0.47)	66.50 (23.88-195.62)
							O	R	≥38.0	45	7	21	227	0.68 (0.60-0.74)	0.97 (0.95-0.99)	22.79 (11.07 -51.79)	0.33 (0.27-0.43)	69.49 (25.96-194.41)
							A	R	≥38.0	43	13	23	221	0.65 (0.56-0.73)	0.94 (0.92-0.97)	11.73 (6.79-20.81)	0.37 (0.29-0.48)	31.78 (14.10-73.04)
2009	Smitz et al.	Belgium	Geriatric unit pts.	>65	100	ITT_A		R	≥37.8	17	1	1	60	0.94 (0.77-1.00)	0.98 (0.93-1.00)	57.61 (11.59-686.19)	0.06 (0.10-0.24)	1020.00 (47.80-139424.45)
						ITT_B		R	≥37.8	16	2	2	59	0.89 (0.70-0.97)	0.97 (0.91-0.99)	27.11 (8.00-115.09)	0.12 (0.03-0.33)	236.00 (24.51-3989.64)
2007	Duberg et al.	Norway		>18	100	ITT		R	≥37.6	23	18	13	46	0.64 (050-0.76)	0.72 (0.64-0.79)	2.27 (1.37-3.60)	0.50 (0.30-0.79)	4.52 (1.74-11.94)
2006	Rajee & Sultana	Australia	Emergency department pts.	18-91	194	ITT		O	≥38.0	6	6	4	178	0.60 (0.30-0.84)	0.97 (0.95-0.98)	18.40 (6.03-43.34)	0.41 (0.16-0.74)	44.50 (8.14-269.56)
2005	Nordås et al.	Norway	Emergency department pts.	18-88	211	ITT_C		R	≥38.0	3	0	19	156	0.14 (0.04-0.14)	1.00 (0.99-1.00)	-	0.86 (0.86-0.97)	-
						ITT_D		R	≥38.0	12	5	10	184	0.55 (0.36-0.68)	0.97 (0.95-0.99)	20.62 (7.63-60.04)	0.47 (0.33-0.67)	44.16 (11.42-183.21)
2002	Varney et al.	USA	Emergency department pts	>60	95	ITT		R	≥38.0	12	0	15	63	0.44 (0.32-0.44)	1.00 (0.95-1.00)	-	0.56 (0.56-0.72)	-
							O	R	≥38.0	12	0	17	66	0.41 (0.30-0.41)	1.00 (0.95-1.00)	-	0.59 (0.59-0.74)	-
2000	Giuliano et al.	USA	Critically ill adults		72 (203)	ITT_E		PA	≥37.0	25	35	40	103	0.39 (0.29-0.49)	0.75 (0.70-0.79)	1.52 (0.95-2.35)	0.82 (0.65-1.02)	1.84 (0.93-3.62)
						ITT_F		PA	≥37.0	41	47	24	91	0.63 (0.52-0.73)	0.66 (0.61-0.71)	1.85 (1.34-2.47)	0.56 (0.39-0.78)	3.31 (1.71-6.42)
							O	PA	≥37.0	40	25	25	113	0.62 (0.51-0.71)	0.82 (0.77-0.86)	3.40 (2.24-5.08)	0.47 (0.34-0.63)	7.23 (3.55-14.84)
2000	Smitz et al.	Belgium	Geriatric unit & ICU pts.	67-91	45	ITT		R	≥37.6	12	4	2	27	0.86 (0.62-0.97)	0.87 (0.76-0.92)	6.64 (2.62-12.62)	0.16 (0.03-0.50)	40.50 (5.27-420.57)
2000	Valle et al.	Norway	Adults inpatients		191	ITT		R	≥38.0	10	0	23	158	0.30 (0.20-0.30)	1.00 (0.98-1.00)	-	0.70 (0.70-0.81)	-
1999	Prentice & Moreland	Canada	Geriatric chronic care unit		30	ITT		O	≥37.5	7	1	5	17	0.58 (0.35-0.66)	0.94 (0.79-1.00)	10.50 (1.62-224.23)	0.44 (0.34-0.83)	23.80 (1.95-661.91)
1997	Christensen et al.	Denmark	Geriatric unit	66-95	121	ITT_A		R	≥38.0	3	7	5	106	0.38 (0.11-0.71)	0.94 (0.92-0.96)	6.05 (1.33-18.27)	0.67 (0.31-0.97)	9.09 (1.37-59.51)
						ITT_B		R	≥37.5	6	30	2	83	0.75 (0.37-0.96)	0.74 (0.71-0.75)	2.83 (1.25-3.81)	0.34 (0.06-0.90)	8.30 (1.40-63.34)
1997	Petersen & Hauge	Norway	Neurosurgery unit pts.	20-78	65 (201)	ITT (Total)		R	≥37.8	21	32	12	136	0.64 (0.47-0.78)	0.81 (0.78-0.84)	3.34 (2.11-4.79)	0.45 (0.27-0.68)	7.44 (3.10-18.08)
						ITT (TNrs)		R	≥37.8	14	20	6	74	0.70 (0.48-0.86)	0.79 (0.74-0.82)	3.29 (1.86-4.86)	0.38 (0.17-0.70)	8.63 (2.65-29.33)
						ITT (UtNrs)		R	≥37.8	7	12	6	62	0.54 (0.28-0.78)	0.84 (0.79-0.88)	3.32 (1.34-6.45)	0.55 (0.25-0.91)	6.03 (1.47-25.39)
1995	Yaron et al.	USA	Emergency department pts.	17-91	100	ITT		R	≥38.5	6	2	4	88	0.60 (0.31-0.76)	0.98 (0.95-1.00)	27.00 (5.78-175.34)	0.41 (0.24-0.73)	66.00 (7.95-730.27)
1995	Skupski et al.	USA	Obstetric population		73	ITT		O	≥38.0	15	1	1	56	0.94 (0.75-0.99)	0.98 (0.93-1.00)	53.44 (10.57-640.64)	0.06 (0.01-0.27)	840.00 (38.95-115667.74)

Note. TP= True positive; FP= False positive; FN= False negative; TN= True negative; SN= Sensitivity; SP= Specificity; PLR= Positive likelihood ratio; NLR= Negative likelihood ratio; DOR= Diagnosis odds ratio; pts.= patients; ITT= Infrared tympanic thermometer; PA= Pulmonaryl artery thermometer; A= Axillary thermometer; R= Rectal thermometer; O= Oral thermometer; ITT_A= ThermoScan PRO 3000; ITT_B= FirstTemp Genius 3000A; ITT_C= Core-Check, model 2090; ITT_D= OtotempLighTouch LTX; ITT_E= FirstTemp Genius II; ITT_F= Tympanic-Pro1; TNrs= Trained Nurses; UtNrs= Untrained Nurses.

3.2 Assessing the Risk of Bias

The results of a quality evaluation for the fifteen final selected studies, utilizing QUADAS-II as the quality assessment tool, suggested that the risk of selection bias was low in almost all of the papers, although two studies (Duberg et al., 2007; Nordås et al., 2005) showed a bias in the domain of patient selection. The following reasons indicated a low risk of selection bias in the studies: 1) the temperatures were measured concurrently or based on a pre-standard to avoid measurement bias, 2) the reference standard was established by theoretical standard, and 3) none of the studies had skipped the process of temperature measurement. Although some studies (Nordås et al., 2005; Smitz et al., 2009; Varney et al., 2002) did not include all patients for analysis, this did not influence the results in the domain of flow and timing.

3.3 Diagnostic Accuracy of Infrared Temperature Measurement

An evaluation of diagnostic accuracy for infrared tympanic temperature measurementwas carried out to compare the results of diagnostic accuracy according to reference standards. Pooled estimates of the sensitivity, specificity, positive and negative diagnostic likelihood ratios, the diagnosis odd rate, and the sROC are summarized in Table 2.

3.3.1 Diagnostic Accuracy of Infrared Tympanic Thermometers

The infrared tympanic temperature readingwas reported in all studies. The sensitivity of the infrared tympanic temperature measurement in adults ranged from 0.14 to 0.94, and specificity was from 0.66 to 1.00.In a meta-analysis, the pooled sensitivity was 0.59(95%, CI 0.55, 0.63) (Figure 2-A) and the pooled specificity was 0.91(95% CI 0.90, 0.92) (Figure 2-B), with high heterogeneity (I^2) among studies 80.1% (X^2=90.40, p<.001), 94.6% (X^2=332.17, p<.001) respectively. The AUC of the SROC was 0.85 and the Q^* value was 0.78(Figure 3-A).

Table 2. Summary results of Meta-analysis

A. The results of diagnostic test accuracy

Thermometer (vs reference standard)		Study No.	Sensitivity (95% CI)	Specificity (95% CI)	Positive likelihood ratio (95% CI)	Negative likelihood ratio (95% CI)	Diagnostic odds ratio (95% CI)
Infrared tympanic	Total	15 (19)	0.59 (0.55, 0.64)	0.91 (0.90, 0.92)	10.51 (5.84, 18.90)	0.45 (0.34, 0.58)	32.70 (14.33, 74.63)
	Rectal	10 (13)	0.58 (0.52, 0.64)	0.93 (0.92, 0.94)	11.55 (5.74, 23.25)	0.45 (0.33, 0.62)	34.44 (14.88, 79.68)
	Oral	3	0.74 (0.57, 0.87)	0.97 (0.94, 0.99)	19.96 (9.20, 43.33)	0.31 (0.12, 0.78)	75.80 (12.61, 455.54)
	Core	2 (3)	0.58 (0.51, 0.66)	0.83 (0.79, 0.86)	3.48 (0.99, 12.29)	0.49 (0.24, 1.02)	9.31 (1.39, 62.29)
Axillary		2	0.75 (0.67, 0.83)	0.97 (0.95, 0.98)	53.61 (0.66, 4385.80)	0.23 (0.08. 0.67)	233.02 (2.16, 25154.70)
Oral		3	0.61 (0.53, 0.68)	0.93 (0.90, 0.95)	12.63 (2.12, 75.15)	0.45 (0.32, 0.64)	29.97 (4.49, 199.82)

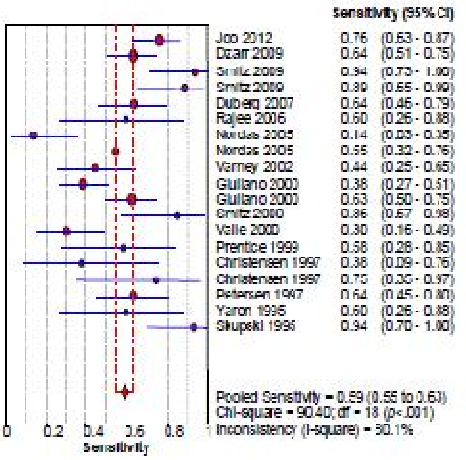
B. The results of summary receiver operating characteristic curve

Thermometer		Study No.	AUC	SE(AUC)	Q*	SE(Q*)
Infrared tympanic	Total	15 (19)	0.85	0.05	0.78	0.05
	Rectal	10 (13)	0.87	0.05	0.80	0.05
	Oral	3	1.00	0.01	0.99	0.02
	Core	2 (3)	0.50	0.28	0.50	0.21
Axillary		2	0.50	0.00	0.50	0.00
Oral		3	0.74	0.15	0.69	0.12

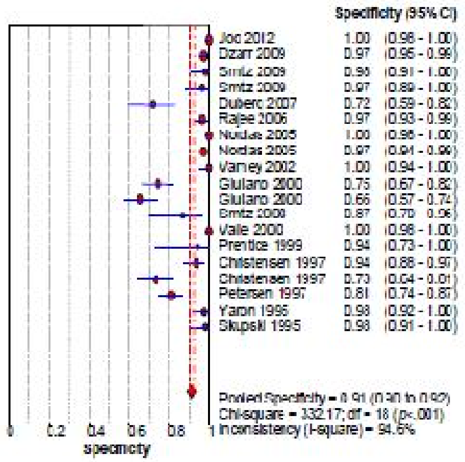
Note. AUC= Area under the curve; SE= Standard error.

Diagnostic accuracy for infrared tympanic temperature readings included three detailed analyses. First, diagnostic accuracy for infrared tympanic temperaturewas compared according to a reference standard. Ten of the studiesused rectal temperature as the "gold standard." In the case of rectal temperature as the reference standard, the sensitivity of the infrared tympanic temperature ranged from 0.14 to 0.89, and specificity was from 0.72 to 1.00. On a meta-analysis, the pooled sensitivity was 0.58(95%, CI 0.53, 0.64) while the pooled specificity was 0.93 (95% CI 0.91, 0.94), the AUC of the SROC was 0.87 and the Q^* value was 0.80.

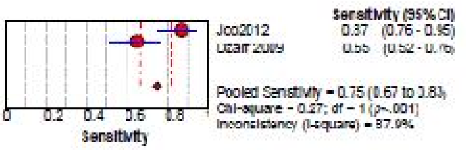
In cases where oral temperature was used as the reference standard (Prentice & Moreland, 1999; Rajee & Sultana, 2006; Skupski et al., 1995) the sensitivity of the infrared tympanic temperature ranged from 0.58 to 0.94, and specificity was from 0.94 to 0.98. In the meta-analysis, the pooled sensitivity was 0.74(95%, CI 0.57, 0.87) and the pooled specificity was 0.97(95% CI 0.94, 0.99), while the AUC of the SROC was 1.00 and the Q^* value was 0.99. In studies in which core (pulmonary artery) temperature was used as the reference standard (Giuliano et al., 2000; Joo & Sohng, 2012) the sensitivity of the infrared tympanic temperature ranged from 0.38 to 0.76, while the specificity was from 0.66 to 1.00. In a meta-analysis, the pooled sensitivity was 0.58(95%, CI 0.51, 0.66) and the pooled specificity was 0.83(95% CI 0.79, 0.86), while the AUC of the SROC was 0.50 and the Q^* value was 0.50.



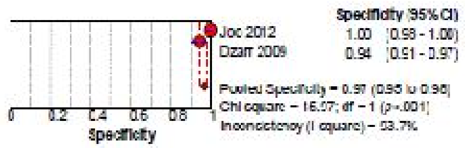
A. Sensitivity of Infra-red tympanic thermometer



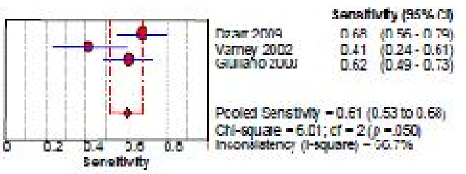
B. Specificity of Infra-red tympanic thermometer



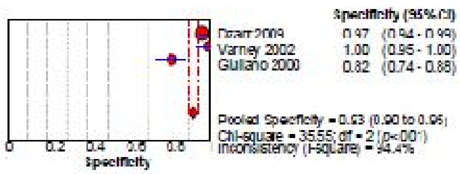
C. Sensitivity of axillary thermometer



D. Specificity of axillary thermometer



E. Sensitivity of oral thermometer



F. Specificity of oral thermometer

Figure 2: Diagnosis Test Accuracy

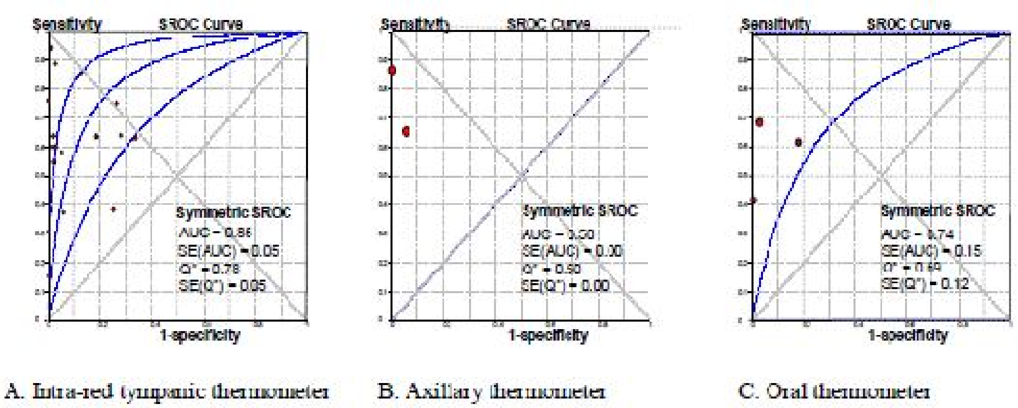


Figure 3: Summary Receiver Operating Characteristic Curve

Moreover, there was a difference in sensitivity between nurses who had received training on the reference standard method of the infrared tympanic thermometer for six months and non-trained nurses. The sensitivity was 0.70 (95% CI 0.48, 0.86) and 0.54 (95% CI 0.28, 0.78) respectively. However, no significant difference in specificity between the two groups was shown. The specificity was found to be 0.79 (95% CI 0.74, 0.82) and 0.84(95% CI 0.79, 0.88), respectively(Petersen & Hauge,1997).

Two out of three studies which compared infrared tympanic thermometersfrom different clinical thermometer manufacturer (Guilano et al., 2000; Nordås et al., 2005; Smitz et al., 2009) demonstrated differences in sensitivity.

3.3.2 Diagnosis Accuracy of Infrared Axillary Thermometers

Infrared axillary temperature was reported in two studies (Dzarr et al., 2009; Joo & Sohng, 2012). The sensitivity of the infrared axillary temperature measured in adults ranged from 0.65 to 0.87, while specificity was from 0.94 to 1.00.According to the meta-analysis, the pooled sensitivity was 0.75(95% CI 0.67, 0.83) (Figure 2-C) and the pooled specificity was 0.97(95% CI 0.95, 0.98) (Figure 2-D), with high heterogeneity (I^2)among studies87.9% (X^2=8.27, p<.001), 93.7% (X^2=15.97, p<.001) respectively. The AUC of the SROC was 0.50 and the Q^* value was 0.50 (Figure 3-B).

3.3.3 Diagnosis Accuracy of Infrared Oral Thermometers

Infrared oral temperature was reported in three studies(Dzarr et al., 2009; Guilano et al., 2000; Varney et al., 2002). The sensitivity of the infrared oral temperature measured in adults ranged from 0.41 to 0.68, and the specificity was from 0.82 to 1.00.In a meta-analysis, the pooled sensitivity was 0.61(95% CI 0.53, 0.68) (Figure 2-E) and the pooled specificity was 0.93(95% CI 0.90, 0.95) (Figure 2-F), with high heterogeneity (I^2)among studies87.9% (X^2=8.27, p<.001), 94.4% (X^2=35.55, p<.001) respectively. The AUC of the SROC was 0.74 and the Q^* value was 0.69 (Figure 3-C).

4. Discussion

Diagnostic accuracy is a test's abilitytodeterminethe existence of disease or the status of a person. According to the reference standard, the existence of disease will be classified as positive or negative. After bisecting subjects into categories of positive and negative, a contingency table can be drawn. Through this method, sensitivity and specificity, positive and negative predictive values, positive and negative diagnostic likelihood ratios, or AUC (Lee & Lee, 2011) can be expressed as a measure of diagnostic accuracy. The present study aimed to examine the diagnostic accuracy of infrared tympanic temperature measurement through a systematic review and meta-analysis after systemic consideration of the various methods of temperature measurement in a total of 1,468 subjects (2,179 distinct measurements) from 15 diagnostic assessment studies, in order to investigate the diagnostic accuracy of infrared tympanic temperature measurement.

Although mathematical standards about sensitivity and specificity exist in diagnosis assessment, the meta-analysis results of infrared tympanic temperature measurement in the present study were interpreted as compared to axillary and oral temperature measurements. Since it is not possible to exclude all factors that affect body temperature, a "gold standard" cannot exist, although this study did assess diagnostic accuracy according to a reference standard.

The results of the present study found that the diagnostic accuracy of infrared tympanic temperature measurement was not less than that of axillary or oral temperature measurements.

The pooled sensitivity of infrared tympanic temperature measurement was 0.59(95%, CI 0.55, 0.63), the pooled specificity was 0.91(95% CI 0.90, 0.92), and the AUC of the SROC was 0.85. This indicates that infrared tympanic temperature can be used to measure the body temperature for adult patients in a clinical setting with a similar level of accuracy as axillary temperature measurements (sROC AUC 0.50) or oral temperature measurements (sROC AUC 0.74). In the case of axillary temperature, the AUC of the sROC was found to be 0.50, which can be interpreted as a meaningless diagnostic test. However, the result of two studies that used the same subjects and that investigated infrared temperature showed similarity within sensitivity and specificity, suggestingproblems with the reference standard.

Subgroup analyses according to reference standard found the following: the diagnostic accuracy of infrared tympanic temperature appeared high compared with that of rectal or oral temperature as the reference standard, whereas the diagnostic accuracy appeared low compared with core(pulmonary artery) temperature as the reference standard. This result suggests that when rectal or oral temperature are used as the reference standard, the infrared tympanic temperature may be the appropriate diagnostic test but when core(pulmonary artery) is the reference standard, it may be not. This findingdiffered from conclusions drawn from studies in which the infrared tympanic temperature reflected the core (pulmonary and bladder) temperature (Jefferies, Weatherall, Young, & Beasley, 2011; Joo & Sohng, 2012). Due to the restrictedamount of research that uses core temperature as a reference standard, the present study has a limitation. Some studies reported that tympanic temperature is underestimated in cases of over 38.3°C and overestimated in cases of lower than 38.3°C (Schmitz, Bair, Falk, & Levine, 1995); the selected studies involved the lower cut-off point for fever, which may have influenced the sensitivity and specificity.

The pooled specificity (the percentage of afebrile adults who were correctly identified as not having a fever) was 0.91 and the pooled sensitivity (the percentage of febrile adults who were correctly identified as having a fever) was 0.59, ranging from 0.14 to 0.94 with the higher deviation. Most papers demonstrated more than moderate levels of sensitivity, and as the study with low sensitivity (Nordås et al., 2005) reported this sensitivity on a specific machine,the possibility of a problem with that machine cannot be excluded.

Variability of a measuring instrument is the most important factor in judging the efficiency of instruments(Spitzer, 2008), and some studies(Park & Park, 2007) have reported that not all thermometers on the market are accurate and that more than half of all infrared tympanic thermometers arebeyond allowable limits. Therefore, it is important to use an accurate and deliberate thermometer to satisfy international standards. The selected papers reviewed in this study included only febrile participants and the small numbers of febrile cases may have affected sensitivity. Additionally, the various cut-off points for fever in the studies, witha lower cut-off point leading to higher sensitivity(Christensen et al., 1998), may have affected the sensitivity and specificity. Therefore, further studies for the diagnostic accuracy of infrared temperature should set the cut-off point for the fever in the reference standard to analyze the accuracy of the infrared temperature measurement. Also since in clinical practice tympanic temperature may have limits in detecting fever, it is appropriate to use more than one type of thermometer and to measure temperature repeatedly.

Heterogeneity among studies appears high in terms of sensitivity and specificity, with specificity in particular appearing as more than 90%. Measuring sites, gender differences, patient age, and ambient temperature may have affected the normal range of body temperature and contributed to this result(Lu & Dai, 2009; Sund-Levander, Forsberg, & Wahren, 2002).

Since tympanic temperature rapidly reflects core temperature even when body temperature is changing, it is likely to become the gold standard for measuring temperature in elderly people (Sund-Levander & Grodzinsky, 2013). However, Lu, Leasure, & Dai(2009) reported that persons of 60 years and older tend to have a lower normal body temperature thanother adults, with rectal temperatures of 0.4°C, tympanic temperatures of 0.2°C, oral temperatures of 0.7°C, and axillary temperatures 0.3°C lower. Cautionshould be used when diagnosing geriatric patients as fever can be easily overlooked due to these patients' lower normal body temperatures. It is therefore recommended that in future studies, patients be divided into categories of adults below and over 60 years of age.

The pooled sensitivity of infrared tympanic temperature measured by trained nurses was 0.70 (95% CI 0.48, 0.86), whereas that measured by non-trained nurse was 0.54 (95% CI 0.28, 0.78).

The advantage of using tympanic thermometers is that they are relatively easy and fast to use, and they provide comparatively accurate readings; however,because an individual'slack of skills may affect the sensitivity and specificity, users do require sufficient training (Spitzer, 2008). For measurements of tympanic membrane temperature, the pinna should be gently pulled backwards and the ear thermometer tip inserted into the external auditory meatus. It is also important to repeat the measurement after two to three minutes because ear wax and positioning could impact tympanic temperature(Sohng et al., 2009; Yoo & Jo, 2009).

Since infrared tympanic thermometers have become convenient and popular to use in clinical settings, the reliability of tympanic temperature measurement needs to be ensured and scientific consensus is required. Therefore, this study is meaningful in its contribution to establish nursing practice guidelines by clarifying the diagnostic accuracy of infrared tympanic temperature measurement through systemic literature review and meta-analysis.However, the study does present some limitations. First, it encompassed only two studiesin which core (pulmonary artery) temperature was used as the reference standard. Secondly, studies were analyzed including both rectal temperature and oral temperature as the reference standard. Although the standard of temperature evaluation involves measuring the core (pulmonary artery) temperature, only patients with a pulmonary catheter could be measured from this site (Joo & Sohng, 2012) requiring other patientsto bemeasured from another site as the reference standard. Therefore, when subjects are not good candidates for a pulmonary catheter, other temperatures sites should be selected as the referral standard(Jefferies et al., 2011). Due to this limitation, further studies are required to measure diagnostic accuracy compared against core (pulmonary artery) temperature as the reference standard.

5. Conclusion

The results of this systematic review of 15 diagnostic assessment studies suggest that infrared tympanic temperature measurement is a useful method to examine accurate temperature in adults. However, an accurate temperature measurement depends onan accurate thermometer utilized by a trained individual. Furthermore, as measuring sites, gender differences, patient age, and ambient temperature may affect body temperature, it is important that temperature readings be taken from the same site consistently for each patient.

It may be necessary to measure temperature in more than one way to detect fever. In future, further research based on scientific evidence through systemic review should be conducted in order to establish nursing practice guidelines regarding temperature measurement and to provide suggestions for clinical practice.

References

Childs, C., Harrison, R.,& Hodkinson, C. (1999). Tympanic membrane temperature as a measure of core temperature. Archives of Disease in Childhood, 80(3), 262-266.

Christensen, P.M., Christensen, V.B. & Matzen, L.E. (1998). Evaluation of tympanic temperature measurements in a geriatric department. Ugeskr Laeger, 160(36), 5175-5177.

Duberg, T., Lundholm, C.,& Holmberg, H.(2007). Tympanic thermometer not an adequate alternative to rectal thermometer. Lakartidningen, 104(19), 1479-1482.

Dzarr, A.A., Kamal, M.,& Baba, A.A. (2009). A comparison between infrared tympanic thermometry, oral and axilla with rectal thermometry in neutropenic adults. European Journal of Oncology Nursing, 13(4), 250-254. http://dx.doi.org/10.1016/j.ejon.2009.03.006

Giuliano, K.K., Giuliano, A.J., Scott, S.S., MacLachlan, E., Pysznik, E.,Elliot, S. et al.(2000). Temperature measurement in critically ill adults: a comparison of tympanic and oral methods. American Journal of Critical Care,9(4), 254-261.

Greiner, M., Pfeiffer, D., & Smith, R.D. (2000). Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. Preventive Veterinary Medicine, 45(1-2), 23-41.

Higgins, J.P., & Thompson, S.G. (2002). Quantifying heterogeneity in a meta-analysis. Statistics in Medicine, 21(11), 1539-1558.

Jefferies, S., Weatherall, M., Young, P., & Beasley, R. (2011). A systematic review of the accuracy of peripheral thermometer in estimating core temperatures among febrile critically ill patients. Critical Care and Resuscitation, 13(3), 194-199.

Jeong, I.S., & Yoo, E.J. (1997). Study on the body temperature measuring time and accuracy and reliability of tympanic thermometer. Journal of Korean Academy of Fundamentals of Nursing, 4(1), 19-30.

Joo,G., & Sohng,K.A. (2012). Accuracy, precision, and validity of fever detection using non-invasive temperature measurement in adult coronary care unit patients with pulmonary catheters. Journal of Korean Academy of Nursing,42(3), 424-433. http://dx.doi.org/10.4040/jkan.2012.42.3.424

Kozier, B., Erb, G., Blais, K., & Wilkinson, J.M. (1997). Fundamentals of nursing: Concepts, process, and practice(5th ed.). Addison-Wesley:Redwood City, CA.

Lee, T.J., & Kim, D.S. (2007). Fever. Korean Journal of Pediatrics, 50(2), 121-126.

Lee, Y. K., & Lee, S. M. (2011). Clinical trials and accuracy of diagnostic tests. Journal of Genetic Medicine, 8(1), 28-34. http://dx.doi.org/10.5734/jgm.2011.8.1.28

Lu, S.H., & Dai, Y.T.(2009). Normal body temperature and the effects of age, sex, ambient temperature and body mass index on normal oral temperature: aprospective, comparative Study. International Journal of Nursing Studies, 46(5), 661-668. http://dx.doi.org/10.1016/j.ijnurstu.2008.11.006.Epub 2009Jan6

Lu, S.H., Leasure, A.R., & Dai, Y.T.(2009). A Systematic review of body temperature variations in older people. Journal of Clinical Nursing, 19(1), 4-16. http://dx.doi.org/10.1111/j.1365-2702.2009.02945.x. Epub 2009 Nov 3

Macaskill, P., Gatsonis, C., Deeks, J.J., Harbord, R.M., & Takwoingi, Y. (2010). Chapter 10: Analysing and presenting results. In J. J. Deeks, P. M. Bossuyt & C. Gatsonis (ed.), Cochrane handbook for systematic reviews of diagnostic test accuracy version 1.0 (pp. 1-47). The Cochrane Collaboration:Oxford, UK.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D.G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. Annals of Internal Medicine, 151(4), 264-269.http://dx.doi.org/10.7326/0003-4819-151-4-200908180-00135

Nordås, T.G., Leiren, S., & Hansen, K.S. (2005). Can tympanic temperature measurement be used in a hospital?Tidsskr Nor Laegeforen, 125(20), 2763-2765.

Park, C.W., & Park, S.N.(2007). Development of standard and improvement of reliability in body temperature measurement. Journal of the Korean Society for Precision Engineering, 24(9), 32-36.

Park, Y.J., Park, S.H., & Kang, C.B.(2013). Systematic review and meta-analyses of diagnostic accuracy of infrared thermometer when identifying fever in children. Journal of Korean Academy Nursing, 43(6), 746-759. http://dx.doi.org/10.4040/jkan.2013.43.6.746

Petersen, M.H., &Hauge, H.N.(1997). Can training improve the results with infrared tympanic thermometers? Acta Anaesthesiologica Scandinavica, 41(8), 1066-1070.

Prentice, D., & Moreland, J. (1999). A comparison of infrared tympanic thermometry with electronic predictive thermometry in a geriatric setting. Geriatric Nursing,20(6), 314-317.

Rajee, M., & Sultana, R.V. (2006). NexTemp thermometer can be used interchangeably with tympanic or mercury thermometers for emergency department use. Emergency Medicine Australasia, 18(3), 245-251.

Schmitz, T., Bair, N., Falk, M., & Levine, C.1995. A comparison of five methods of temperature measurement in febrile intensive care patients. American Journal of Critical Care, 4(4), 286-292.

Skupski, D.W., Sonnenblick, A.L., Wagner, W.E., & Chervenak, F.A. (1995). An evaluation of tympanic thermometry in an obstetric population. Journal of Maternal-Fetal Investigation, 5(3), 152-154.

Smitz, S., Giagoultsis, T., Dewe, W., & Albert, A. (2000). Comparison of rectal and infrared tympanic temperatures in older hospital inpatients. Journal of American Geriatric Society,48(1), 63-66.

Smitz, S., Van de Winckel, A., & Smitz, M.F. (2009). Reliability of infrared tympanic thermometry in the prediction of rectal temperature in older inpatients. Journal of Clinical Nursing, 18(3), 451-456.

Sohng, K.Y., Park, H.S., Hong, Y.H., Yun, E.J., Lee, K.Y.,& Cho, B.H. (2009). Fundamentals of nursing(2nded.). Soomoonsa:Seoul.

Spitzer, O.P. (2008). Comparing tympanic temperatures in both tympanics to oral temperature in the critically ill adult. Dimensions of Critical Care Nursing, 27(1), 24-29.

Sund-Levander, M., Forsberg, C., & Wahren, L.K.(2002). Normal oral, rectal, tympanic and axillary body temperature in adult men and women: asystematic literature review. Scandinavian Journal of Caring Science, 16(2), 122-128.

Sund-Levander, M., & Grodzinsky, E.(2013). Assessment of body temperature measurement options. British Journal of Nursing, 22(15), 880-888.

Valle, P.C., Kildahl-Andersen, O.,& Steinvoll, K. (1999). A comparative study of infrared tympanic thermometry and rectal mercury thermometry. Scandinavian Journal of Infectious Disease, 31(1), 105-106.

Van Staaij, B.K., Rovers, M.M., Schilder, A.G., & Hoes, A.W. (2003). Accuracy and feasibility of daily infrared tympanic membrane temperature measurements in the identification of fever in children. International Journal of Pediatric Otorhinolaryngology, 67(10), 1091-1097.

Varney, S.M., Manthey, D.E., Culpepper, V.E., & Creedon, J.F.J.(2002). A comparison of oral, tympanic, and rectal temperature measurement in the elderly. Journal of Emergency Medicine, 22(2), 153-157.

Walter, S.D. (2002). Properties of the summary receiver operating characteristic (SROC) curve for diagnostic test data. Statistics in Medicine, 21(9), 1237-1256.

Wells, N., King, J., Hedstrom, C., & Youngkins, J. (1995). Does tympanic temperature measure up? MCN. The American Journal of Maternal Child Nursing, 20(2), 95-100.

Whiting, P.F., Rutjes, A.W., Westwood, M.E., Mallett, S., Deeks, J.J., &Reitsma, J. et al. (2011). QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. Annals of Internal Medicine, 155(8), 529-536.
http://dx.doi.org/10.7326/0003-4819-155-8-201110180-00009

Yaron, M., Lowenstein, S.R., & Koziol-McLain, J.(1995). Measuring the accuracy of the infrared tympanic thermometer: correlation does not signify agreement. The Journal of Emergency Medicine, 13(5), 617-621.

Yoo, J.H., & Jo, H.S.(2009). Comparison of Tympanic and Axillary temperatures. Journal of Korean Academy of Fundamentals of Nursing, 16(2), 162-170.

Yun, K.W., & Lim, I.S. (2005). A study for accuracy and usefulness of tympanic membrane and forehead thermometers. Korean Journal of Pediatrics, 48(8), 820-825.