Lee *et al. BMC Infectious Diseases* (2025) 25:31
https://doi.org/10.1186/s12879-024-10332-0

BMC Infectious Diseases

SYSTEMATIC REVIEW

Open Access

The impact of age on comparative diagnostic accuracy of temporal artery thermometers and non-contact infrared thermometers for fever detection: a systematic review and meta-analysis

Kwo-Chen Lee[1,2], Yun-Ping Lin[1,2], Ya-Ling Tzeng[1,2], Wen-Chun Liao[1,2], Chyi Lo[1,2], Pei-Yun Chen[3] and Shu-Hua Lu[1,2*]

Abstract

Background Non-invasive temporal artery thermometers (TATs) and non-contact infrared thermometers (NCITs) are increasingly used in community settings to measure body temperature. Existing research predominantly focuses on pediatric populations, yet the accuracy and precision of TATs and NCITs for fever screening across age groups remain unclear. This study aims to assess age-related differences in the diagnostic accuracy of TATs and NCITs for fever detection.

Methods A systematic review and meta-analysis were conducted, sourcing data from PubMed, MEDLINE, CINAHL, EMBASE, Cochrane Library, ProQuest, and Web of Science. Prospective studies comparing TATs and NCITs against body temperature measurement methods were included. Two independent researchers extracted data, and study quality was assessed with the QUADAS-2 tool. Pooled estimates of sensitivity, specificity, and the hierarchical summary area under the receiver operating characteristic (ROC) curves were calculated using STATA version 17.

Results This meta-analysis included 34 studies with 28,996 participants, of whom 5,358 were febrile. For TATs, 22 studies with 9,894 readings yielded a pooled sensitivity of 0.59 (95% CI: 0.40–0.76) and specificity of 0.91 (95% CI: 0.83–0.96). Sensitivity was higher at fever thresholds > 38 °C (0.71, 95% CI: 0.60–0.80), and higher in children (0.77, 95% CI: 0.66–0.85) than in adults (0.48, 95% CI: 0.30–0.67). Similar sensitivities were observed between rectal and other standards (0.70, 95% CI: 0.59–0.80 vs. 0.70, 95% CI: 0.41–0.89). For NCITs, 16 studies with 14,234 readings yielded a pooled sensitivity of 0.70 (95% CI: 0.54–0.82) and specificity of 0.94 (95% CI: 0.90–0.97). Sensitivity improved at fever thresholds > 38 °C (from 0.70 to 0.75, 95% CI: 0.55–0.88) and was higher in children compared to the overall estimate (0.79 vs. 0.70, 95% CI: 0.62–0.90). Comparable sensitivities were noted between axillary and other standards (0.73, 95% CI: 0.30–0.94 vs. 0.75, 95% CI: 0.49–0.90).

*Correspondence:
Shu-Hua Lu
shuhua@mail.cmu.edu.tw
Full list of author information is available at the end of the article

Conclusions TATs and NCITs show variable diagnostic accuracy across age groups, with higher sensitivity in children and at elevated fever thresholds. This variability underscores the importance of age-specific use of these thermometers and highlights the need for further research to optimize diagnostic performance across populations.

Keywords Temporal artery thermometers, Non-contact infrared thermometers, Diagnostic accuracy, Sensitivity, Specificity

Introduction

Fever is a critical clinical indicator of infection, making accurate body temperature measurement essential for early diagnosis and effective treatment monitoring [1]. Ensuring precise temperature readings that closely reflect core body temperature is challenging, especially during outbreaks of infectious diseases where rapid, non-invasive methods are paramount [2].

Temporal Artery Thermometers (TATs) and Non-Contact Infrared Thermometers (NCITs) are non-invasive tools designed to detect fever swiftly, which is particularly valuable in various clinical settings. TATs function by scanning the temporal artery, offering ease of use and proximity to core temperature measurements. NCITs provide contactless measurements, which is advantageous for reducing the risk of cross-infection [3, 4]. Despite these practical benefits, the accuracy of these devices is influenced by factors such as age of the patient, measurement site, and external conditions, resulting in conflicting evaluations of their diagnostic reliability across different patient demographics [5, 6].

While existing research largely centers on pediatric populations, the diagnostic accuracy of TATs and NCITs in adults remains insufficiently explored. Some studies indicate that TATs align closely with core temperature, whereas others warn of potential inaccuracies in febrile patients [7, 8]. Similarly, NCITs, commonly used in community and clinical settings, have shown variable sensitivity rates in adults, raising concerns about their reliability across various age groups [9, 10]. These inconsistencies highlight the necessity of age-specific evaluations to determine the diagnostic performance of these thermometers effectively.

This systematic review and meta-analysis aims to assess the diagnostic accuracy of TATs and NCITs across different age groups. The primary objective is to determine the impact of age on the comparative diagnostic performance of these thermometers for fever detection, providing crucial insights into their clinical utility across diverse age demographics.

Methods

Design

This systematic review was conducted in accordance with the methodological standards outlined in the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy [11]. The review follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses of Diagnostic Test Accuracy Studies (PRISMA-DTA) guidelines [12] as specified in the PRISMA-DTA checklist. The study protocol was registered with the International Prospective Register of Systematic Reviews (PROSPERO, registration number CRD 42022363943).

Database search

We conducted a comprehensive search of relevant databases to identify eligible articles published between January 2000 and October 2022. The databases searched included PubMed, MEDLINE, CINAHL, EMBASE, Cochrane Library, ProQuest, and Web of Science.

Search strategy

The search strategy was developed using a combination of relevant keywords and filters. Key search terms included "temporal artery temperature," "non-contact infrared thermometers," "fever," "temperature," and age-specific terms such as "adult," "children," and "elderly." The full search strategy, including all specific combinations, is available in the Supplementary Materials.

The search results were managed using Endnote 21 bibliographic software. After removing duplicates, the full texts of all potentially relevant articles were retrieved and reviewed to ensure they met the inclusion criteria.

Study eligibility

We included studies measuring body temperature with TATs or NCITs, considered as the index tests. These were compared to reference standards for body temperature measurement, which included rectal thermometry (widely regarded as the clinical gold standard), as well as tympanic, oral, and axillary thermometers. The reference standard varied by study design and population, with measurements conducted in hospital settings or community out-patient clinics. We included prospective diagnostic accuracy studies, randomized controlled trials, controlled clinical trials, and observational studies.

Studies lacking explicit statistical values for sensitivity and specificity, not reporting the number of participants or the incidence of fever, and studies not published in English or Chinese were excluded.

Study selection

Two independent reviewers (SHL and KCL) conducted an initial screening of titles and abstracts to assess the potential eligibility of identified studies. Any discrepancies in the selection process were resolved through discussion, and in cases where consensus could not be reached, a third reviewer (YPL) was consulted to make the final decision. Subsequently, the full texts of all potentially relevant studies were meticulously reviewed to determine their eligibility for inclusion in the analysis. The quality of the included studies was independently assessed by the two primary reviewers (SHL and KCL) using predefined criteria.

Data extraction and qualitative synthesis

Data extraction was independently carried out by two reviewers (SHL and KCL). The following information was systematically extracted from each study: 1. Study details: Author(s) and year of publication. 2. Participant demographics: Age of participants, study site, total number of participants, and the number of individuals with confirmed fever cases. 3. Index test: Thermometer brand and model, as well as the site of temperature measurement. 4. Reference standard: Thermometer brand and model, measurement site, and the defined fever threshold used in the study. 5. Diagnostic accuracy measures: Studies were identified as diagnostic accuracy studies if they reported key metrics, including sensitivity (the proportion of febrile individuals correctly identified as having a fever) and specificity (the proportion of non-febrile individuals correctly identified as not having a fever) [13].

Methodological quality of included studies

The methodological quality of each included study was evaluated using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool [14], which assesses studies across four key domains: patient selection, index test, reference standard, and flow and timing. Any disagreements between the two primary reviewers (SHL and KCL) regarding the quality assessments were discussed, and a consensus was reached through the involvement of a third reviewer when necessary (YPL).

Statistical analysis

Statistical analysis involved extracting or calculating data for 2 × 2 contingency tables (true positives, false negatives, false positives, and true negatives) from individual studies (see Supplementary Tables 1 and 2). Any ambiguities or missing data were clarified by directly contacting the authors of the respective studies. Heterogeneity was assessed by visually inspecting the width of the Summary Receiver Operating Characteristics (HSROC) curve and the prediction region. A wider HSROC curve is indicative of the presence of substantial heterogeneity [11]. Publication bias was assessed using Deeks' funnel plot asymmetry test, conducted in STATA (version 17, Stata Corp, College Station, TX, USA) through the midas module. Sensitivity analyses were performed to explore the impact of variables such as patient age, fever threshold, and type of reference standard on diagnostic outcomes.

Results

Search outcome

A total of 1,315 potentially relevant citations were identified, of which 34 met the inclusion criteria and were included in both qualitative and quantitative synthesis (Fig. 1). The literature search and study selection process are illustrated in a PRISMA diagram [47].

Characteristics of included studies

1. TATs

A total of 22 studies included in the qualitative synthesis were published between 2002 and 2022, spanning 15 countries. The majority of these studies were conducted in the United States ($n = 11$) [15–18, 22, 23, 26, 27, 29, 31, 48], followed by Canada [21], Australia [20], the United Kingdom [19], Denmark [30, 32], Turkey [33], Germany [24], Belgium [8], and Nigeria [25, 46].

Most of these studies reported on one index test device (per study) except in one study [23], where three devices were compared. A total of five devices were investigated, namely Exergen [8, 15, 18, 19, 21, 22, 24–30], SensorTouch [16], TAT-5000 [33, 46, 48], Beurer FT 60 infrared forehead thermometer digital contact forehead thermometer, and the Chicco Thermo Touch Plus-contact forehead thermometer [23].

Most of these studies reported on one reference standard device per study. However, in three studies [16, 31, 48], two devices were compared. The most common reference standard was rectal temperature ($n = 19$) [8, 15, 17, 18, 22, 24–32, 48], followed by oral temperature ($n = 7$) [19, 23, 31, 34, 48], bladder temperature ($n = 3$) [16, 20, 21], axillary temperature ($n = 1$), and pulmonary artery temperature catheter ($n = 1$) [16].

Most of the studies reported on one defined fever threshold per study, except in one study [16], where three thresholds were compared, and one [18], where two thresholds were compared. The defined fever threshold range was 37.6–39 °C, with 38 °C being the most common ($n=21$) [8, 15, 17, 18, 22, 23, 25–32, 34, 48], followed by 37.8 °C ($n=3$) [16, 20], 38.5 °C ($n=2$) [15, 18], 39 °C ($n=1$) [15] 38.3 °C ($n=1$) [21], 37.6 °C ($n=1$) [19], and 37.9 °C ($n=1$) [24].

Most study settings were in the emergency room ($n=10$) [17, 22, 24, 27, 29, 31, 32, 48], followed by pediatric clinics ($n=7$) [8, 15, 19, 25, 26, 28, 30], intensive care units ($n=3$) [18, 20, 21], operating room ($n=2$) [16, 20], and hospitals ($n=1$) [23]. Only one study [22] reported the ambient temperature.

A total of 9,894 readings from 6,919 patients were obtained in these 22 studies. The age groups of most research subjects represented children ($n=14$) [8, 15–19, 21, 22, 25, 26, 28–30, 48] (age range: 4 days–18 years), followed by adults ($n=7$) [16, 20, 24, 27, 29, 31, 32, 46]. There were three age groups that included both children and adults ($n=3$) [16, 23, 33] (age range: 4 days–87 years).

In the studies where TATs were used as the index test, the reported prevalence of fever ranged from 4% to 96.8%. The sensitivity of TATs across these studies varied between 0 and 0.967, while specificity ranged from 0.465 to 1.000 (Table 1).

2. **NCITs**

A total of 16 studies included in the qualitative synthesis were published between 2004 and 2022, representing 15 countries, i.e., the United States [23], Australia [9], the United Kingdom [43], Italy [39], Turkey [40, 41], France [37], the Netherlands [38], Nigeria [46], Korea [44], China [42, 45], Hong Kong [36], and Taiwan [35]. The detailed characteristics of the included studies are presented in Table 2.

Most of these studies reported on a single index test device per study, except for one study [43], which compared two devices. In total, 16 NCITs were evaluated across these studies, representing 12 distinct device models. The investigated thermometer brands or models included Raynger MX-4 [37], Thermofocus [22, 35, 38, 39, 43], Thermoflash LX-26 [40, 48], DT 8806 [41], Standard-8812 [36], JXB-170 [45], HuBDIC Thermofinder S2 FBF-710 [44], Cocoon-NC 9900 [9], Firhealth forehead thermometer [43], VeraTemp [33], and TriTemp, [46].

All of these studies reported on one reference test device (per study). The most common reference indicators were axillary temperature ($n=8$) [33, 39, 41, 43, 49], followed by tympanic temperature ($n=7$) [35–37, 42, 44], rectal temperature ($n=2$) [22, 38], oral temperature ($n=2$) [23, 46], and forehead temperature ($n=1$) [9].

Most of these studies reported on one defined fever threshold, except in the study by Hausfater [37], where three thresholds were compared. The defined fever thresholds ranged from 37.3–38.5 °C, with 38 °C being the most common ($n=11$) [22, 23, 33, 34, 36–40, 43, 44], followed by 37.5 °C ($n=3$) [9, 35, 37], 37.3 °C ($n=3$) [42, 49], and 38.5 °C ($n=1$) [37].

The study settings included: emergency room ($n=7$) [22, 37, 39, 41, 42, 49], pediatric ward ($n=4$) [36, 38, 40, 44], out-patient department ($n=2$) [43, 46], ward ($n=1$) [9], and outdoor ($n=1$) [35]. The ambient temperature was not provided in more than half of the studies. Only five studies [22, 37, 39–41] reported the ambient temperature.

A total of 14,234 readings from 8,877 patients were collected in these 16 studies. The age groups of most research subjects represented children ($n=9$) [22, 36, 38–41, 43, 44] (age range: 4 day–18 years), followed by adults ($n=3$) [9, 46, 49]. One of the articles categorized adults as over 16 years old [9]. There were three research age groups that included both children and adults ($n=2$) [23, 37] (age range: 4 days–103 years).

In the studies where NCITs were used as the index test, the reported prevalence of fever ranged from 3% to 57.6%. The sensitivity of NCITs across these studies varied between 0.125 and 0.970, while specificity ranged from 0.600 to 1.000 (Table 2).

Methodological quality of included studies

The results of the quality assessment for the included studies ($n=34$) are summarized in Fig. 2. Overall, five of the 34 studies were found to have a high risk of bias in at least one of the four domains of the QUADAS-2 tool [16, 31, 35–37], while two studies exhibited a high risk of bias in two domains [19, 23].

Quantitative data synthesis

1. **TATs**

Diagnostic accuracy of TATs

A total of 22 studies on TATs, comprising 9,894 individual readings, were included in our analysis. The pooled sensitivity and specificity for TATs, irrespective of age distribution, fever threshold, and reference site, were 0.59 (95% CI: 0.40–0.76) and 0.91 (95% CI: 0.83–0.96) (Fig. 3A), respectively. A summary receiver operating characteristic (ROC) curve indicated an overall diagnostic accuracy of 0.87 (95% CI: 0.84–0.90) (Fig. 3B). No evidence of publication bias was detected, as indicated by Deeks' funnel plot asymmetry test ($p=0.23$) (Fig. 3C).

Table 1 Characteristics of the included studies with temporal artery thermometers (TATs) as the index tests

No	Author/Year of publication	Setting/Ambient temperature	Sample characteristics	Index thermometer/ Device type	Reference standard (device)/ Device type	Fever threshold (°C)	Fever N (%)	AUC (%)	Sensitivity	Specificity
01	Siberry, 2002 [15]	Pediatric	275 children Age < 24 months mean 11.2 months	Home model Exergen	Rectal/Digital	38	47 (17.5)	*NA*	0.93	0.46
	Siberry, 2002 [15]	Pediatric	275 children Age < 24 months, mean 11.2 months	Home model Exergen	Rectal/Digital	38.5	30 (11)	*NA*	0.96	0.83
	Siberry, 2002 [15]	Pediatric	275 children Age < 24 months, mean 11.2 months	Home model Exergen	Rectal/Digital	39	21 (8)	*NA*	0.95	0.95
02	Suleman, 2002 [16]	OR	30 adults Age 36–83 years	SensorTouch	Pulmonary artery	37.8	15 (50)	*NA*	0	100
	Suleman, 2002 [16]	OR	26 children Age 9 days–13 years, mean 3 ± 4 years	SensorTouch	Bladder temperature	37.8	16 (61.5)	*NA*	0.84	0.83
03	Callanan, 2003 [17]	ER	187 children Age < 3 months	SensorTouch	Rectal/Digital (Welch Allyn SureTemp)	38	23 (12.2)	*NA*	0.83	0.86
04	Hebbar, 2005 [18]	PICU	44 children	Exergen	Rectal/Digital	38	19 (43.1)	*NA*	0.44	0.89
	Hebbar, 2005 [18]	PICU	44 children Age mean 11.5 months	Exergen	Rectal/Digital	38.5	19 (43.1)	*NA*	0.40	0.95
05	Ei-Radhl, 2007 [19]	Pediatric	34 children Age 4–16 years	Exergen TAT 5000	Oral/(Welch Allyn SureTemp)	37.6	15 (56)	*NA*	0.19	0.98
06	Kimberge, 2007 [20]	OR (n = 35) ICU (n = 35)	70 adults Age range: 18–80 years, mean 49 years	Exergen TAT 5000	Bladder temperature	> = 37.8	55 (19.6)	*NA*	0.72	0.97
07	Stelfox, 2010 [21]	ICU	760 readings from 14 children	Exergen TAT 500	Bladder temperature	> = 38.3	736 (96.8)	0.80	0.29	0.99
08	Teran, 2011 [22]	ER, Inpatient unit 24–28 °C	434 children	Exergen TAT 2000	Rectal/Glass mercury	> = 38	167 (38.4)	*NA*	0.91	0.99

Table 1 (continued)

No	Author/Year of publication	Setting/Ambient temperature	Sample characteristics	Index thermometer/ Device type	Reference standard (device)/ Device type	Fever threshold (°C)	Fever N (%)	AUC (%)	Sensitivity	Specificity
09	Hamilton, 2013 [23]	Hospital	171 subjects Age 4 days–87 years, > 18 years (n = 19)	Chicco Thermo Touch Plus -contact forehead thermometer	Oral/Welch Allyn SureTemp Plus model 692	> = 38	64 (37.4)	*NA*	0.72	0.95
		Hospital	171 subjects Age 4 days–87 years, > 18 years (n = 19)	Beurer FT 60 infrared forehead thermometer digital contact forehead thermometer	Oral/Welch Allyn SureTemp Plus model 692	> = 38	64 (37.4)	*NA*	0.45	0.95
		Hospital	171 subjects Age 4 days–87 years, > 18 years (n = 19)	Exergen TAT 2000C	Oral/Welch Allyn SureTemp Plus model 692	> = 38	64 (37.4)	*NA*	0.54	0.95
10	Hamilton, 2013 [23]	ER	n = 205 New bone ~ < 18 years, mean 66 months	TAT 500	< 5 years: rectal > 5 years: oral	> = 38	93(46)	*NA*	0.72	0.96
11	Singler, 2013 [24]	R	427 older adults Age > = 75 years	Exergen Temporal Scanner™	Rectal/IVAC® TEMP PLUS II® Model 2080 Thermometer	37.9	105 (24.5)	0.64	0.38	0.91
12	Allegaert, 2014 [8]	Pediatric	294 children Age < 17 years median 3.2 years	Exerogen	Rectal/Filac 3000	38	22 (7.48)	*NA*	0.41	0.98
13	Odinaka, 2014 [25]	Pediatric	156 children Age < 59 months mean 10.8 ± 13.6 months	Exergen TAT 2000C	Rectal/Mercury	38	79 (50.64)	*NA*	0.83	0.88
14	Moore, 2015 [26]	Pediatric	239 children Age 3 months–4 years, 1.5 ± 0.8 years	Exergen	Rectal/Digital	38	98 (41)	*NA*	0.56	0.93
15	Bijur, 2016 [27]	ER	987 adults	Exergen TAT 500	Rectal/Electronic	> = 38	284 (28.7)	*NA*	0.71	0.92
16	Forrest, 2017 [28]	ediatric	85 children Age < 36 months, mean 12 ± 8 months	Exergen TAT 5000	Rectal/Welch Allyn	38	26 (30.5)	*NA*	0.61	0.93
17	Brosinski, 2018 [29]	ER	126 children Age < 3 years	Exergen TAT 5000	Rectal/Welch Allyn SureTemp Plus	38	57 (45.2)	*NA*	0.90	0.85

Table 1 (continued)

No	Author/Year of publication	Setting/Ambient temperature	Sample characteristics	Index thermometer/ Device type	Reference standard (device)/ Device type	Fever threshold (°C)	Fever N (%)	AUC (%)	Sensitivity	Specificity
		ER	125 older adults Age > 65 years	Exergen TAT 5000	Rectal/Welch Allyn SureTemp Plus	38	47 (37.6)	*NA*	0.74	0.94
18	Mogensen, 2018 [30]	Pediatric ward	n = 965 children Age < 18 median 24 months	Exergen TAT 500	Rectal/OMRON	> = 38	377 (39)	0.93	0.81	0.90
19	Haimovich, 2020 [31]	ER	1293 adults	TAT/Not reported	Rectal/NA	> = 38 (100.4°F)	444 (34.4)	*NA*	0.27	0.98
		ER	1293 adults	TAT/Not reported	Oral/NA	> = 38 (100.4°F)	55 (4.3)	*NA*	0.23	0.99
20	Nygarrd, 2020 [32]	ER	381 Adults Age > = 18 years	TAT/Not reported	Rectal/Digital	> = 38	36 (9.4)	*NA*	0.66	95
21	Erdem, 2021 [33]	ER	151 subjects Age 2–18 years	TAT 5000-IR	Axillary/Digital	> = 38	67 (44.3)	*NA*	0.93	0.94
22	Ravi, 2022 [34]	OPD	200 adults	Exergen TAT 5000	Oral digital/Welch Allyn SureTemp Plus 690	> = 38	8 (4)	0.87	0.88	0.88

Abbreviations : *ER* Emergency Room, *OR* Operating Room, *OPD* Out-Patient Departments, *ICU* Intensive Care Unit, *NA* Not Available

Table 2 Characteristics of the included studies with handheld non-contact infrared thermometers (NCITs) as the index tests

No	Author/Year of publication	Setting/Ambient temperature	Sample characteristics	Index thermometer/ Device type	Reference standard (device)/Device type	Fever threshold (°C)	Fever N (%)	AUC	Sensitivity	Specificity
1	Liu, 2004 [35]	Outdoor	500 subjects	Thermofocus	Tympanic/Digital (Welch Allyn SureTemp)	37.5	110 (22)	*NA*	0.17	0.98
02	Ng, 2005 [36]	Pediatric ward	1000 children Age: median 2 years	Standard ST-8812	Tympanic/ FirstTemp	38	123 (12.3)	0.86	0.89	0.75
03	Hausfater, 2008 [37]	ER 24.8 °C	2026 subjects Age range: 6–103 years	Rayngner MX	Tympanic/Braun Pro 4000	> =37.5	62 (3)	0.93	0.76	0.66
	Hausfater, 2008 [37]	ER 24.8 °C	As above	Rayngner MX	Tympanic/Braun Pro 4000	> =38	62 (3)	0.87	0.82	0.77
	Hausfater, 2008 [37]	ER 24.8 °C	As above	Rayngner MX	Tympanic/Braun Pro 4000	> =38.5	62 (3)	0.79	0.82	0.90
04	Paes, 2010 [38]	Pediatric ward	100 subjects Age range: 2 weeks–18 years	Thermofocus	Rectal/Terumo C402	> =38	25 (25)	*NA*	0.64	0.96
05	Chiappini, 2011 [39]	Pediatric ER 24.15 °C	251 children Age > 1 month, median age 4.5 years	Thermofocus	Axially/Mercury	> =38	53 (21.1)	0.96	0.89	0.90
06	Teran, 2011 [22]	ER, inpatient unit 24–28 °C	434 children Age < 48 months mean14.6 ± 10.7 months	Thermofocus	Rectal/Glass mercury	> =38	163 (38.4)	*NA*	0.97	0.97
07	Apa, 2013 [40]	Pediatric 24–26 °C	1639 children Mean age: 54 months	ThermoFlash LX-26	Axially/Microlife MT 3001	> =38	175 (10.6)	0.96	0.94	0.90
08	Hamilton, 2013 [23]	Hospital	171 subjects Age range: 4 days–87 years > 18 years (n = 19)	Visiomed SAS Thermoflash LX-26 noncontact forehead thermometer	Oral/Welch Allyn SureTemp Plus model 692	> =38	67 (37.4)	*NA*	0.44	0.95
09	Atas, 2018 [41]	ER 28 °C	319 subjects Age range: > 1 month–18 years, median 30 months	DT 8806	Axially/Beurer FT09	> =37.5	184 (57.6)	0.79	0.88	0.60
10	Chen, 2020 [42]	ER	261 adults	Not reported	Tympanic/Braun thermoscan PRO 6000	> =37.3	44 (17)	0.81	0.93	0.60
11	Hayward, 2020 [43]	OPD	401 children Age 0–5 years	Thermofocus 0800	Axillary/	> =38 °C	16 (4)	*NA*	0.66	0.98
	Hayward, 2020 [43]	OPD	401 children Age 0–5 years	Firhealth forehead thermometer	Axillary/Welch Allyn SureTemp	> =38 °C	16 (4)	*NA*	0.12	0.99
12	Erdem, 2021 [33]	ER	151 subjects Age 2–18 years	VeraTemp	Axillary/Digital	> =38	67 (44.3)	*NA*	0.69	1.00
13	Khan, 2021 [9]	Ward	265 adults Age 16 years	Cocoon-NC9900	TAT/Exergen TAT 5000	> 37.5	7 (2.6)	0.67	0.16	0.99
14	Kim, 2022 [44]	Pediatric ward 25–27 °C	255 children	Hubdic Thermofinder S2, HFS-710	Tympanic/Braun ThermoScan 7, IRT 6520	> =38	40 (15.6)	*NA*	0.81	0.98

Table 2 (continued)

No	Author/Year of publication	Setting/Ambient temperature	Sample characteristics	Index thermometer/ Device type	Reference standard (device)/Device type	Fever threshold (°C)	Fever N (%)	AUC	Sensitivity	Specificity
15	Lai, 2022 [45]	ER/OPD 14–29 °C	235 adults Mean age 29.6 years 669 children Mean age 3.71 years	JXB-170	Axillary/Mercury	> = 37.3	44.1 (45.5)	*NA*	0.49	0.97
	Lai, 2022 [45]	ER/OPD 14–29 °C	As above	JXB-170	Axillary/Mercury	> = 37.3	441 (45.5)	*NA*	0.47	0.97
16	Ravi, 2022 [46]	OPD 31.5 °C	200 adults	TriTemp, Trimedika	Oral digital/Welch Allyn SureTemp Plus 690	> = 38	8 (4)	0.62	0.13	0.96

Abbreviations: ER Emergency Room, *OR* Operating Rooms, *OPD* Out-Patient Departments, *ICU* Intensive Care Unit, *NA* Not Available

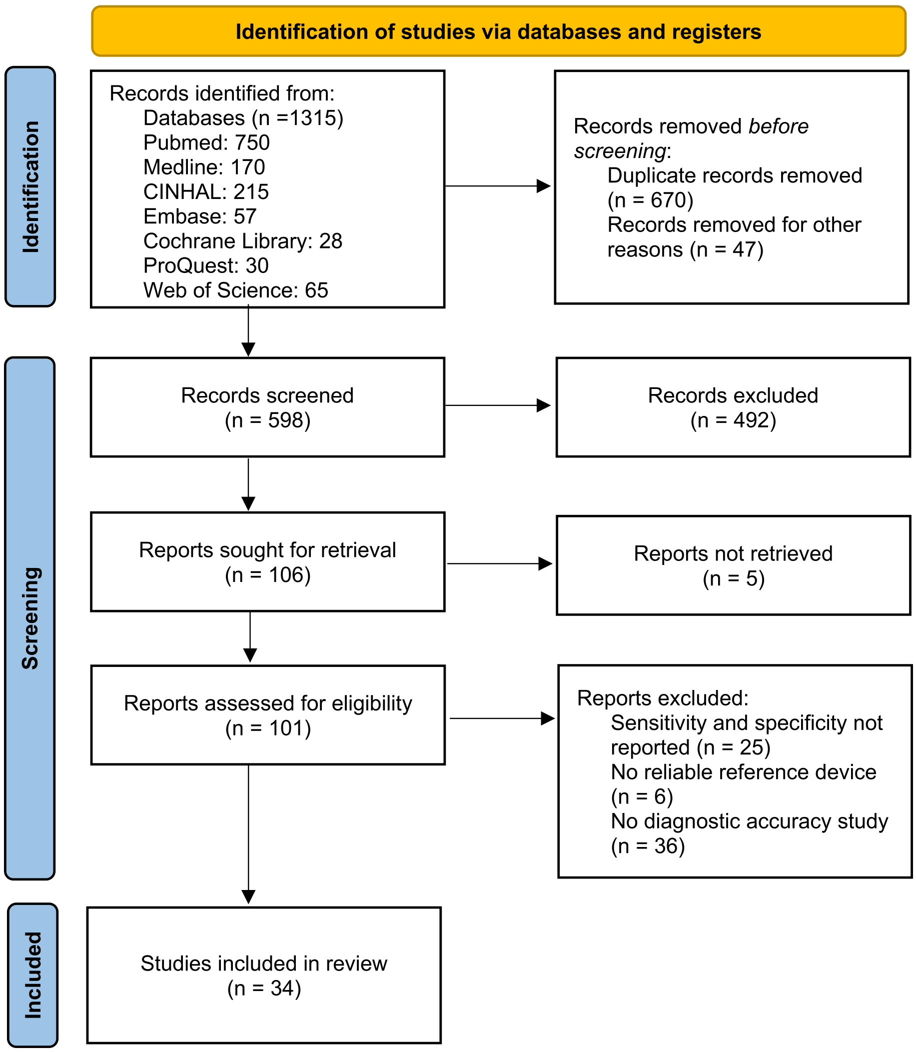


Fig. 1 Preferred reporting items for systematic reviews and meta-analysis (PRISMA) flow diagram of the study selection process

Heterogeneity

The Spearman correlation coefficient for TATs was − 0.501 ($p = 0.025$), indicating the presence of a threshold effect. A wider HSROC curve was observed in Fig. 3B, suggesting heterogeneity among TAT studies. Subgroup analyses were conducted to identify potential sources of this heterogeneity.

Sensitivity analysis

The results are summarized in Table 3, with detailed forest plots available in Supplementary Fig. 1.

1–1. Age group

Subgroup analyses revealed that sensitivity was higher at fever thresholds exceeding 38 °C, and significantly greater in children (0.77, 95% CI: 0.66–0.85) compared to adults (0.48, 95% CI: 0.30–0.67) (Table 3; Supplementary Fig. 1A, B). These findings suggest that TATs demonstrate better performance in the pediatric population, with sensitivity improving when adult populations are excluded from the analysis.

1–2. Fever threshold

When the analysis was restricted to studies using a fever threshold of greater than 38 °C (excluding thresholds below 38 °C), the pooled sensitivity increased from 0.59 (95% CI: 0.40–0.76) to 0.71 (95% CI: 0.60–0.80) (Table 3) (Supplementary Fig. 1C). This indicates that TATs perform more accurately at higher fever thresholds.

1–3. Reference standard

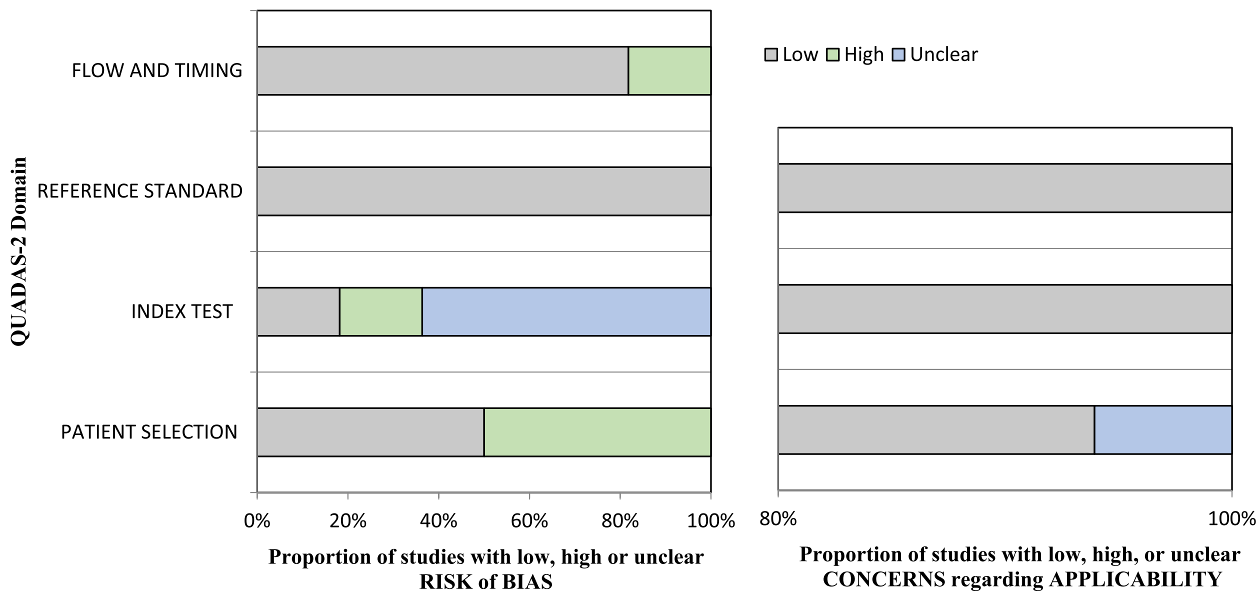


Fig. 2 Quality assessment of included studies with QUADAS-2: summary for risk of bias and applicability concerns

The most common reference standard for TATs was rectal temperature (n = 14). Owing to the limited number of studies using oral, axillary, or tympanic temperatures as reference standards, these three were combined into a single group for analysis. The results showed no difference in pooled sensitivity between rectal temperature and the combined reference standards (0.70, 95% CI: 0.59–0.80 vs. 0.70, 95% CI: 0.41–0.89) (Table 3) (Supplementary Fig. 1D, E).

2. **NCITs**

Diagnostic accuracy of NCITs

A total of 16 studies on NCITs, comprising 14,234 individual readings, were included in our analysis. The pooled sensitivity and specificity for NCITs, regardless of age distribution, fever threshold, and reference site, were 0.70 (95% CI: 0.54–0.82) and 0.94 (95% CI: 0.90–0.97) (Fig. 4A), respectively. A summary ROC curve indicated an overall diagnostic accuracy of 0.93 (95% CI: 0.90–0.95) (Fig. 4B). No evidence of publication bias was detected, as indicated by Deeks' funnel plot asymmetry test (p = 0.7) (Fig. 4C).

Heterogeneity

The Spearman correlation coefficient for NCITs was −0.572 (p = 0.052), suggesting the absence of a significant threshold effect. A wider HSROC curve was observed in Fig. 4B, indicating the presence of heterogeneity. Consequently, further subgroup analyses were conducted to identify potential sources of this heterogeneity.

Sensitivity analysis

The results are summarized in Table 4, with detailed forest plots available in Supplementary Fig. 2.

2–1. Age group

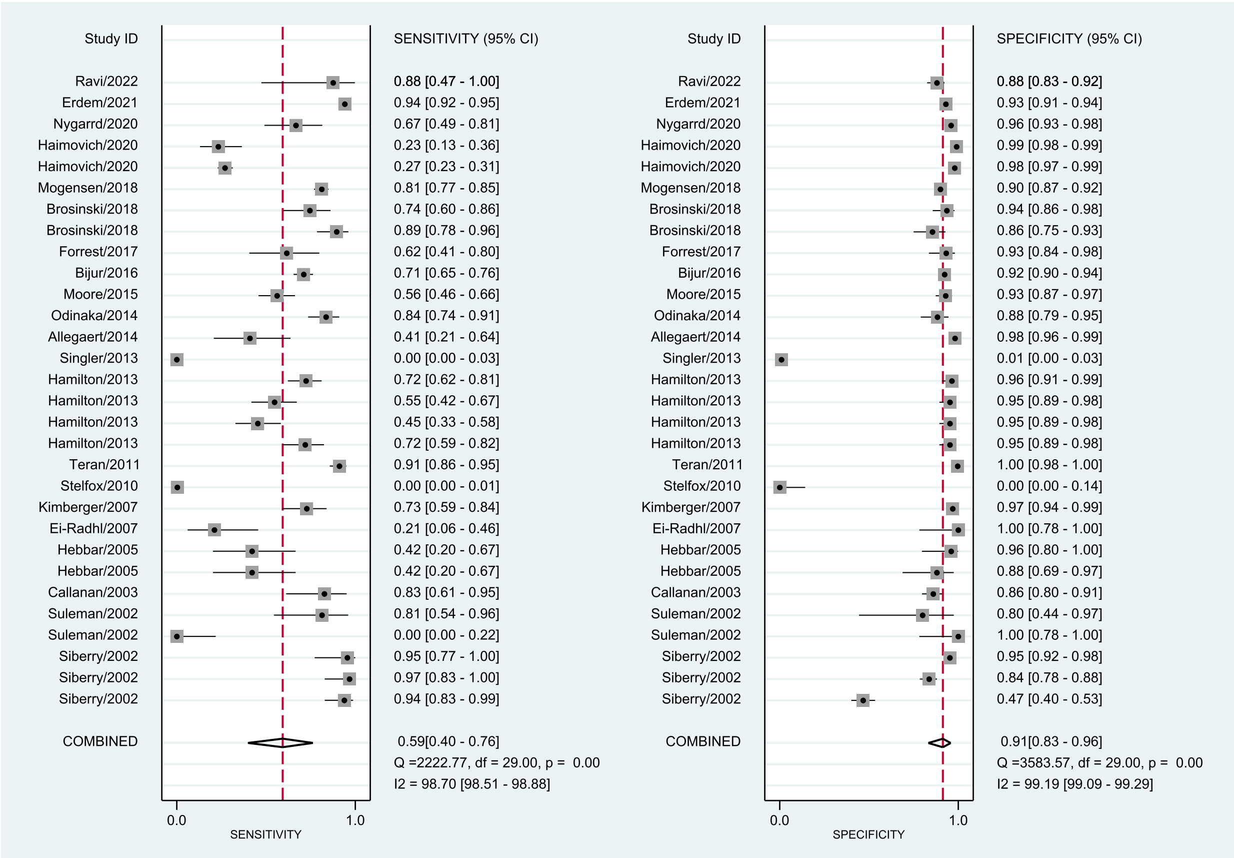
As there was only one adult study with a fever threshold above 38 °C, it was excluded from the analysis. Subsequent analysis, limited to the pediatric population, revealed an increase in the pooled sensitivity from 0.70 (95% CI: 0.54–0.82) to 0.79 (95% CI: 0.62–0.90) (Table 4; Supplementary Fig. 2A). This finding suggests that NCITs exhibit better diagnostic performance when used exclusively in children.

2–2. Fever threshold

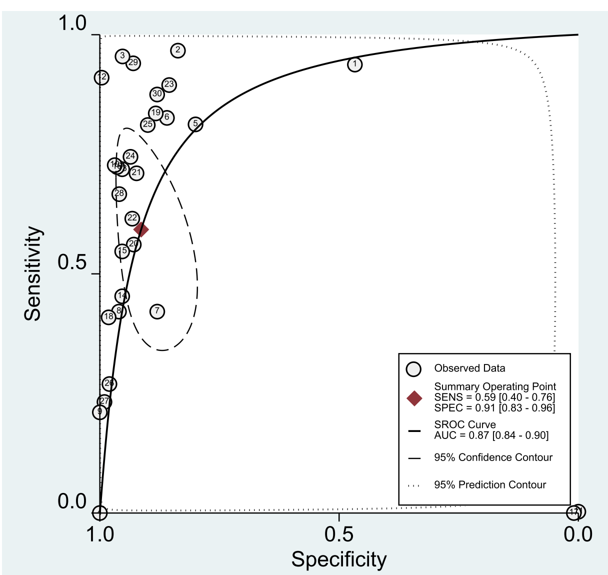
Separate analyses were performed for fever thresholds above and below 38 °C. The pooled sensitivity for fever thresholds >38 °C was 0.75 (95% CI: 0.55–0.88), while for thresholds <38 °C, the pooled sensitivity was 0.59 (95% CI: 0.33–0.81) (Table 4) (Supplementary Fig. 2B, C). These results indicate that NCITs perform more accurately at higher fever thresholds.

2–3. Reference standard

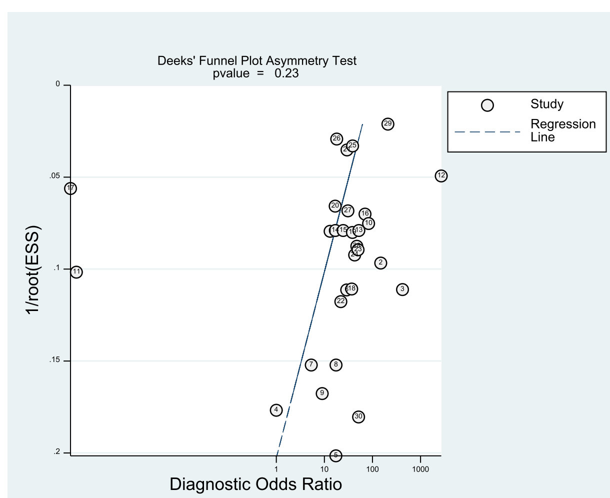
The most common reference standard for NCITs was axillary temperature (n = 5). Due to the limited number of studies using rectal, oral, and tympanic



(A) Forest plot showing pooled sensitivity and specificity values
(Fever threshold: 37.6–39 °C; Reference standard: rectal, oral, bladder, axillary, and pulmonary artery temperature)



(B) Hierarchical summary receiver operating characteristic (HSROC) curves



(C) Funnel plot depicting publication bias

Fig. 3 Quantitative analysis for the overall diagnostic accuracy of temporal artery thermometers (TATs) for the detection of fever using the forehead as the site of measurement

Table 3 Pooled sensitivity and specificity of temporal artery thermometers (TATs) in different subgroups

	Study n/Analysis n	Pooled sensitivity (95% CI)	Pooled specificity (95% CI)
Total	22 studies/30 analyses	0.59 (0.40–0.76)	0.91 (0.83–0.96)
Age group			
Adult	5 studies/6 analyses	0.48 (0.30–0.67)	0.96 (0.93–0.98)
Child	13 studies/15 analyses	0.77 (0.66–0.85)	0.94 (0.90–0.97)
Fever threshold			
> 38 °C	18 studies/25 analyse	0.71 (0.60–0.80)	0.94 (0.91–0.96)
< 38 °C	4 studies/5 analyses	0.09 (0.00–0.81)	0.90 (0.16–1.00)
Reference standard			
Rectal	14 studies/15 analyses	0.70 (0.59–0.80)	0.94 (0.92–0.96)
Oral, TM, Ax	3 studies/6 analyses	0.70 (0.41–0.89)	0.96 (0.92–0.98)

Abbreviations: TM tympanic membrane, *Ax* Axillary

temperatures as reference standards, these were combined into a single group for analysis. The pooled sensitivity was similar between rectal temperatures and the combined reference standards (0.73, 95% CI: 0.30–0.94 vs. 0.75, 95% CI: 0.49–0.90) (Table 4) (Supplementary Fig. 2 D, E), suggesting comparable diagnostic accuracy across different reference standards.

Discussion

This study provides a comprehensive evaluation of the diagnostic accuracy of TATs and NCITs for fever detection. Our pooled analysis revealed moderate sensitivity for both TATs (0.59, 95% CI: 0.40–0.76) and NCITs (0.70, 95% CI: 0.54–0.82), with high specificity for both (0.91, 95% CI: 0.83–0.96 and 0.94, 95% CI: 0.90–0.97, respectively). Notably, sensitivity improved for both types of thermometers when the fever threshold was set above 38 °C, with TATs at 0.71 (95% CI: 0.60–0.80) and NCITs at 0.75 (95% CI: 0.55–0.88).

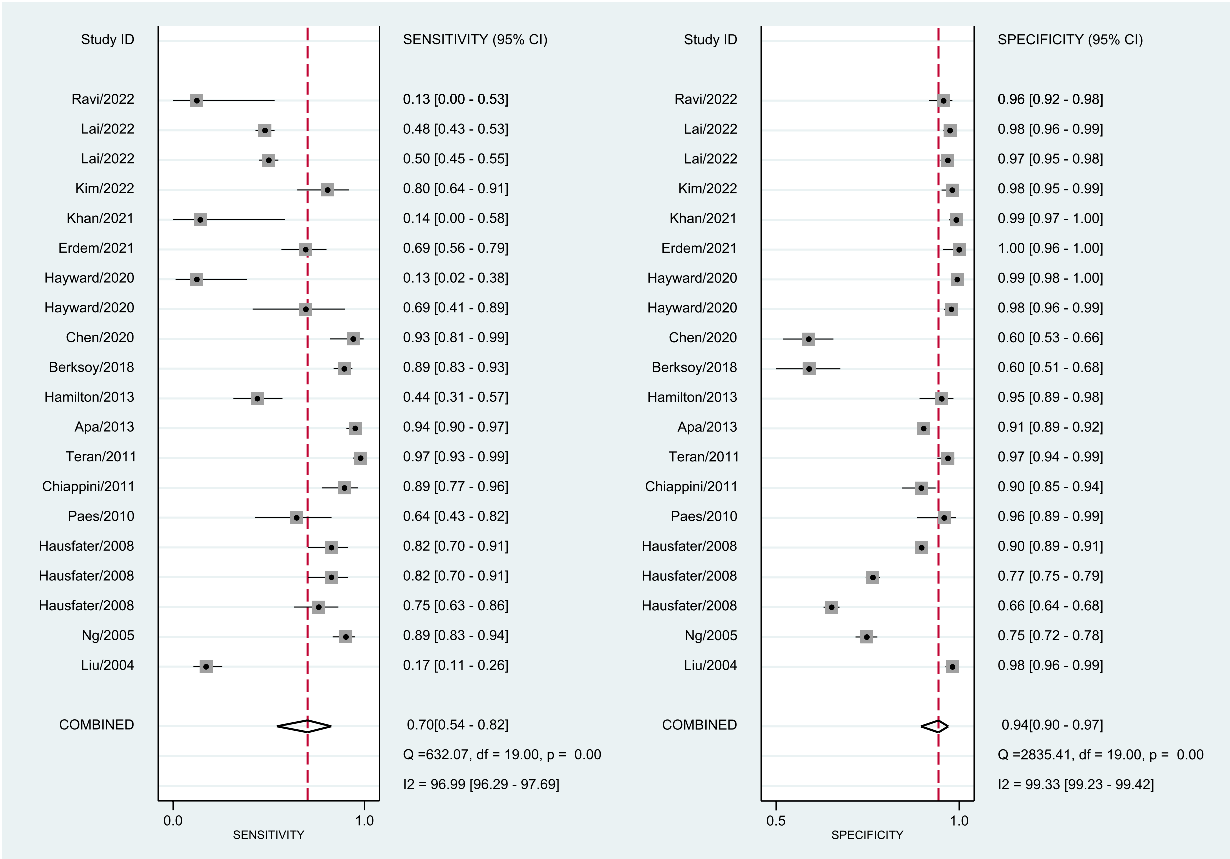
Subgroup analyses revealed that both types of thermometers performed better in pediatric populations compared to adults. For TATs, the pooled sensitivity in children (0.77, 95% CI: 0.66–0.85) was significantly higher than in adults (0.48, 95% CI: 0.30–0.67). Similarly, NCITs exhibited improved diagnostic performance in children, with pooled sensitivity increasing from the overall estimate of 0.70 (95% CI: 0.54–0.82) to 0.79 (95% CI: 0.62–0.90). These findings suggest that patient age plays a critical role in diagnostic accuracy, with age likely contributing as a primary source of heterogeneity in the meta-analysis.

The studies included in this review spanned a wide range of age groups, but few specifically targeted adult or elderly populations, particularly for NCITs, where only one study evaluated a fever threshold above 38 °C in adults [34]. The lower sensitivities observed in studies involving adults may be attributed to physiological differences, such as thicker skin and deeper blood vessels in adults, which reduce the accuracy of infrared thermometry [18, 50]. Additionally, elderly individuals often exhibit diminished skin surface temperatures and attenuated fever responses, further impacting diagnostic sensitivity [51, 52], Other factors such as sweat sebum and makeup may also contribute to falsely low readings, raising concerns about the potential for false negatives and missed febrile cases [3].

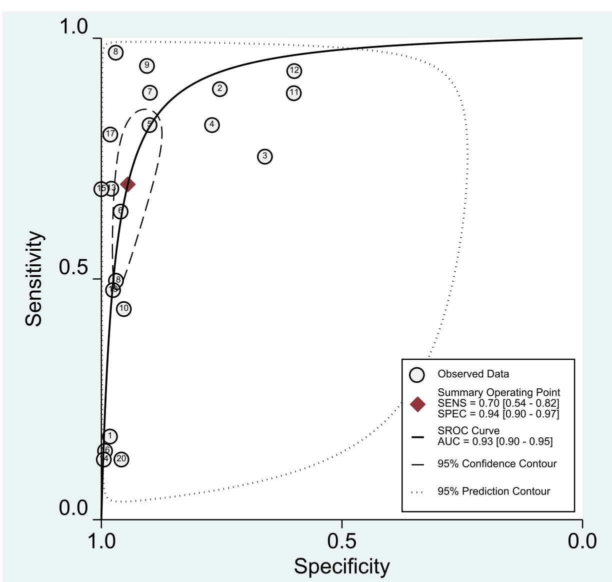
This is the first study to systematically compare the diagnostic accuracy of TATs and NCITs for fever detection. Despite both using infrared technology to measure forehead temperature, the contact-based nature of TATs and the non-contact method of NCITs necessitate separate analyses. Interestingly, our results showed that the pooled sensitivity of NCITs was higher than that of TATs (0.70, 95% CI: 0.54–0.82 vs. 0.59, 95% CI: 0.40–0.76), contrary to initial expectations. However, when restricted to fever thresholds above 38 °C, TATs and NCITs displayed similar sensitivity (0.71, 95% CI: 0.60–0.80 vs. 0.75, 95% CI: 0.55–0.88).

A potential explanation for this discrepancy is the use of different reference standards: TATs predominantly reference rectal temperatures, which are closer to core body temperature, whereas NCITs typically reference axillary temperatures, which reflect surface temperature. These differences in reference standards, despite a common fever threshold, may explain why NCITs performed better in our analysis. The inherent heterogeneity across temperature measurement sites and fever thresholds presents challenges in drawing direct comparisons.

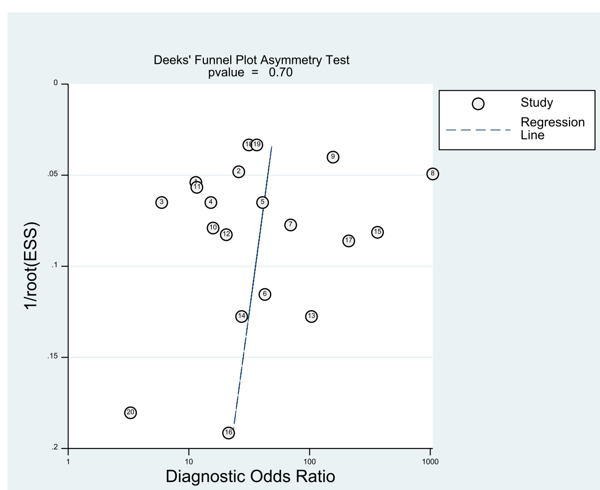
Other sources of heterogeneity include the variety of thermometer models used (five types of TATs and seven types of NCITs), diverse reference standards (primarily rectal for TATs and axillary for NCITs), and varying fever thresholds across studies (ranging from 37.3 °C to 39 °C).



(A) Forest plot showing pooled sensitivity and specificity values
(Fever threshold: 37.3–38.5 °C; Reference standard: tympanic, rectal, oral, and forehead temperature)



(B) Hierarchical summary receiver operating characteristic (HSROC) curves



(C) Funnel plot depicting publication bias

Fig. 4 Quantitative analysis for the overall diagnostic accuracy of non-contact infrared thermometers (NCITs) for the detection of fever using the forehead as the site of measurement

Table 4 Pooled sensitivity and specificity of non-contact infrared thermometers (NCITs) in different subgroups

	Study n/Analysis n	Pooled sensitivity (95% CI)	Pooled specificity (95% CI)
Total	16 studies/20 analyses	0.70 (0.54–0.82)	0.94 (0.90–0.97)
Age group			
Adult	1 study/1 analysis	NA	NA
Child	11 studies/12 analyses	0.79 (0.62–0.90)	0.95 (0.89–0.97)
Fever threshold			
> 38 °C	11 studies/13 analyse	0.75 (0.55–0.88)	0.95 (0.90–0.97)
< 38 °C	6 studies/7 analyses	0.59 (0.33–0.81)	0.92 (0.76–0.98)
Reference standard			
Ax	5 studies/5 analyses	0.73 (0.30–0.94)	0.97 (0.89–0.99)
Rectal Oral, TM	7 studies/7 analyses	0.75 (0.49–0.90)	0.94 (0.87–0.97)

Abbreviations: TM tympanic membrane, *Ax* Axillary, *NA* Not Available

Additionally, ambient temperature has been identified as a significant factor affecting the accuracy of infrared thermometry [53].

Limitations

This study has several limitations. First, variability in reference standards—such as rectal, tympanic, oral, and axillary measurements—introduces potential heterogeneity in diagnostic accuracy. Differences in how each method correlates with core body temperature may affect the comparability of TATs and NCITs. Additionally, the high heterogeneity of research settings and limited data on adult populations present further challenges. Most studies focus primarily on pediatric groups, leading to an insufficient understanding of these devices' accuracy in adults and the elderly. Given the increasing reliance on TATs and NCITs for temperature screening, especially during emerging infectious diseases, future research should prioritize age-stratified studies to assess these devices across a broader demographic spectrum.

Conclusions

This study suggests that NCITs might be more sensitive than TATs for fever detection, although confidence intervals were overlapping and both devices demonstrate comparable sensitivities when measuring temperatures above 38 °C. The diagnostic performance of both types of thermometer is superior in children, with marked declines in sensitivity observed in adult populations. This reduced sensitivity in adults could result in up to 25–30% of febrile cases being undetected. Further research, focused on improving the accuracy of these devices in adult populations, is crucial for optimizing public health strategies, particularly for pandemic preparedness and response.

Abbreviations

TATs	Temporal Artery Thermometers
NCITs	Non-contact Infrared Thermometers
PRISMA-DTA	The Preferred Reporting Items for Systematic Reviews and Meta-Analyses of Diagnostic Test Accuracy Studies
HSROC	Hierarchical Summary Receiver Operating Characteristics (HSROC) curves
ROC	Receiver Operating Characteristics
QUADAS-2	Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12879-024-10332-0.

Supplementary Material 1
Supplementary Material 2
Supplementary Material 3
Supplementary Material 4
Supplementary Material 5

Acknowledgements

We would like to express our sincere gratitude to China Medical University, Taiwan, for providing financial support for this research under grant number CMU110-TC-06. Without their generous assistance, this study would not have been possible.

Authors' contributions

SHL, KCL, and YPL were involved in the conception and design of the study. SHL, KCL, YPL, YLT, WCL, CL, and PYC were discussed inclusion and exclusion criteria. SHL, YPL, YLT, and KCL searched databases, screened articles. SHL, and KCL extracted data independently extracted relevant data, assessed the quality of evidence. SHL and YPL performed meta-analysis, and KCL supervised the research. SHL, KCL, YPL, YLT, WCL, CL, and PYC were involved in the writing, drafting, and editing the final document for publication. All authors have read and agreed with the final version of the manuscript.

Funding

This study was supported by the China Medical University, Taiwan (No. CMU110-TC-06).

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author based on reasonable request.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]School of Nursing, China Medical University, 100, Sec. 1, Jingmao Rd, Taichung 406040, Taiwan. [2]Department of Nursing, China Medical University Hospital, Taichung, Taiwan. [3]Jenteh Junior College of Medicine, Nursing and Management, Miaoli, Taiwan.

Received: 31 July 2024 Accepted: 9 December 2024
Published online: 06 January 2025

References
1. Fitzwater J, Johnstone C, Schippers M, Cordoza M, Norman B. A Comparison of Oral, Axillary, and Temporal Artery Temperature Measuring Devices in Adult Acute Care. Medsurg Nurs. 2019;28(1):35–41.
2. Lawson L, Bridges EJ, Ballou I, Eraker R, Greco S, Shively J, Sochulak V. Accuracy and precision of noninvasive temperature measurement in adult intensive care patients. Am J Crit Care. 2007;16(5):485–96.
3. Pompei F, Pompei M. Non-invasive temporal artery thermometry: Physics, physiology, and clinical accuracy. Proc SPIE. 2004;5404:61–7.
4. Sener S, Karcioglu O, Eken C, Yaylaci S, Ozsarac M. Agreement between axillary, tympanic, and mid-forehead body temperature measurements in adult emergency department patients. Eur J Emerg Med. 2012;19(4):252–6.
5. Aragón-Vargas L. Limitations of temporal (forehead) temperature readings as a screening method for covid-19. Pensar en Movimiento: Revista de Ciencias del Ejercicio y la Salud. 2020;18:e42291.1-10.
6. El-Radhi AS. Infrared thermometers for assessing fever in children: the ThermoScan PRO 4000 ear thermometer is more reliable than the Temporal Scanner TAT-500. Evid Based Nurs. 2014;17(4):115.
7. Wolfson M, Granstrom P, Pomarico B, Reimanis C. Accuracy and precision of temporal artery thermometers in febrile patients. Medsurg Nurs. 2013;22(5):297–302.
8. Allegaert K, Casteels K, van Gorp I, Bogaert G. Tympanic, infrared skin, and temporal artery scan thermometers compared with rectal measurement in children: a real-life assessment. Curr Ther Res Clin Exp. 2014;76:34–8.
9. Khan S, Saultry B, Adams S, Kouzani AZ, Decker K, Digby R, Bucknall T. Comparative accuracy testing of non-contact infrared thermometers and temporal artery thermometers in an adult hospital setting. Am J Infect Control. 2021;49(5):597–602.
10. Tay MR, Low YL, Zhao X, Cook AR, Lee VJ. Comparison of Infrared Thermal Detection Systems for mass fever screening in a tropical healthcare setting. Public Health. 2015;129(11):1471–8.
11. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy. Version 2.0 (updated July 2023). Cochrane, 2023. Available from https://training.cochrane.org/handbook-diagnostic-test-accuracy/current.
12. McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, Clifford T, Cohen JF, Deeks JJ, Gatsonis C, Hooft L, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. JAMA. 2018;319(4):388–96.
13. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, Lijmer JG, Moher D, Rennie D, de Vet HC, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ. 2015;351: h5527.
14. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529–36.
15. Siberry GK, Diener-West M, Schappell E, Karron RA. Comparison of temple temperatures with rectal temperatures in children under two years of age. Clin Pediatr (Phila). 2002;41(6):405–14.
16. Suleman MI, Doufas AG, Akça O, Ducharme M, Sessler DI. Insufficiency in a new temporal-artery thermometer for adult and pediatric patients. Anesth Analg. 2002;95(1):67–71.
17. Callanan D. Detecting fever in young infants: reliability of perceived, pacifier, and temporal artery temperatures in infants younger than 3 months of age. Pediatr Emerg Care. 2003;19(4):240–3.
18. Hebbar K, Fortenberry JD, Rogers K, Merritt R, Easley K. Comparison of temporal artery thermometer to standard temperature measurements in pediatric intensive care unit patients. Pediatr Crit Care Med. 2005;6(5):557–61.
19. El-Radhi AS, Patel SP. Temperature measurement in children with cancer: an evaluation. Br J Nurs. 2007;16(21):1313–6.
20. Kimberger O, Cohen D, Illievich U, Lenhardt R. Temporal artery versus bladder thermometry during perioperative and intensive care unit monitoring. Anesth Analg. 2007;105(4):1042–7.
21. Stelfox HT, Straus SE, Ghali WA, Conly J, Laupland K, Lewin A. Temporal Artery versus Bladder Thermometry during Adult Medical-Surgical Intensive Care Monitoring: An Observational Study. BMC Anesthesiol. 2010;10:13.
22. Teran CG, Torrez-Llanos J, Teran-Miranda TE, Balderrama C, Shah NS, Villarroel P. Clinical accuracy of a non-contact infrared skin thermometer in paediatric practice. Child Care Health Dev. 2012;38(4):471–6.
23. Hamilton PA, Kasbekar RS, Monro R. Clinical performance of infrared consumer-grade thermometers. J Nurs Meas. 2013;21(2):166–77.
24. Singler K, Bertsch T, Heppner HJ, Kob R, Hammer K, Biber R, Sieber CC, Christ M. Diagnostic accuracy of three different methods of temperature measurement in acutely ill geriatric patients. Age Ageing. 2013;42(6):740–6.
25. Odinaka KK, Edelu BO, Nwolisa CE, Amamilo IB, Okolo SN. Temporal artery thermometry in children younger than 5 years: a comparison with rectal thermometry. Pediatr Emerg Care. 2014;30(12):867–70.
26. Moore AH, Carrigan JD, Solomon DM, Tart RC. Temporal Artery Thermometry to Detect Pediatric Fever. Clin Nurs Res. 2015;24(5):556–63.
27. Bijur PE, Shah PD, Esses D. Temperature measurement in the adult emergency department: oral, tympanic membrane and temporal artery temperatures versus rectal temperature. Emerg Med J. 2016;33(12):843–7.
28. Forrest AJ, Juliano ML, Conley SP, Cronyn PD, McGlynn A, Auten JD. Temporal artery and axillary thermometry comparison with rectal thermometry in children presenting to the ED. Am J Emerg Med. 2017;35(12):1855–8.
29. Brosinski C, Valdez S, Riddell A, Riffenburgh RH. Comparison of Temporal Artery Versus Rectal Temperature in Emergency Department Patients Who Are Unable to Participate in Oral Temperature Assessment. J Emerg Nurs. 2018;44(1):57–63.
30. Mogensen CB, Wittenhoff L, Fruerhøj G, Hansen S. Forehead or ear temperature measurement cannot replace rectal measurements, except for screening purposes. BMC Pediatr. 2018;18(1):15.
31. Haimovich AD, Taylor RA, Krumholz HM, Venkatesh AK. Performance of Temporal Artery Temperature Measurement in Ruling Out Fever: Implications for COVID-19 Screening. J Gen Intern Med. 2020;35(11):3398–400.
32. Nygaard H, Maschmann C, Ekmann A. Comparison between temporal and rectal temperature measurement. Dan Med J. 2020;67(11):A04200270.
33. Erdem N, Demirdağ TB, Tezer H, Yayla BCC, Aksakal FNB, Tapısız A, Derinöz O, Okur A, Pınarlı FG, Koçak Ü, et al. The comparison and diagnostic accuracy of different types of thermometers. Turk J Pediatr. 2021;63(3):434–42.
34. Ravi N, Vithyananthan M, Saidu A. Are all thermometers equal? A study of three infrared thermometers to detect fever in an African outpatient clinic. PeerJ. 2022;10: e13283.
35. Liu CC, Chang RE, Chang WC. Limitations of forehead infrared body temperature detection for fever screening for severe acute respiratory syndrome. Infect Control Hosp Epidemiol. 2004;25(12):1109–11.
36. Ng DK, Chan CH, Lee RS, Leung LC. Non-contact infrared thermometry temperature measurement for screening fever in children. Ann Trop Paediatr. 2005;25(4):267–75.
37. Hausfater P, Zhao Y, Defrenne S, Bonnet P, Riou B. Cutaneous infrared thermometry for detecting febrile patients. Emerg Infect Dis. 2008;14(8):1255–8.

38. Paes BF, Vermeulen K, Brohet RM, van der Ploeg T, de Winter JP. Accuracy of tympanic and infrared skin thermometers in children. Arch Dis Child. 2010;95(12):974–8.
39. Chiappini E, Sollai S, Longhi R, Morandini L, Laghi A, Osio CE, Persiani M, Lonati S, Picchi R, Bonsignori F, et al. Performance of non-contact infrared thermometer for detecting febrile children in hospital and ambulatory settings. J Clin Nurs. 2011;20(9–10):1311–8.
40. Apa H, Gözmen S, Bayram N, Çatkoğlu A, Devrim F, Karaarslan U, Günay İ, Ünal N, Devrim İ. Clinical accuracy of tympanic thermometer and non-contact infrared skin thermometer in pediatric practice: an alternative for axillary digital thermometer. Pediatr Emerg Care. 2013;29(9):992–7.
41. Ataş Berksoy E, Bağ Ö, Yazici S, Çelik T. Use of noncontact infrared thermography to measure temperature in children in a triage room. Medicine (Baltimore). 2018;97(5): e9737.
42. Chen G, Xie J, Dai G, Zheng P, Hu X, Lu H, Xu L, Chen X, Chen X. Validity of the Use of Wrist and Forehead Temperatures in Screening the General Population for COVID-19: A Prospective Real-World Study. Iran J Public Health. 2020;49(Suppl 1):57–66.
43. Hayward G, Verbakel JY, Ismail FA, Edwards G, Wang K, Fleming S, Holtman GA, Glogowska M, Morris E, Curtis K, et al. Non-contact infrared versus axillary and tympanic thermometers in children attending primary care: a mixed-methods study of accuracy and acceptability. Br J Gen Pract. 2020;70(693):e236–44.
44. Kim YM, Jang MR, Moon JR, Park G, An YJ, Seo JM. Clinical accuracy of non-contact forehead infrared thermometer measurement in children. Children (Basel). 2022;9(9):1389.
45. Lai F, Li X, Wang Q, Luo Y, Wang X, Huang X, Zhang J, Peng J, Wang Q, Fan L, et al. Reliability of Non-Contact Infrared Thermometers for Fever Screening Under COVID-19. Risk Manag Healthc Policy. 2022;15:447–56.
46. Ravi N, Mathura V, Saidu A. Are all thermometers equal? A study of three infrared thermometers to detect fever in an African outpatient clinic. PeerJ. 2022;10: e13283.
47. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372: n71.
48. Hamilton PA, Marcos LS, Secic M. Performance of infrared ear and forehead thermometers: a comparative study in 205 febrile and afebrile children. J Clin Nurs. 2013;22(17–18):2509–18.
49. Lai F, Li X, Liu T, Wang X, Wang Q, Chen S, Wei S, Xiong Y, Hou Q, Zeng X, et al. Optimal diagnostic fever thresholds using non-contact infrared thermometers under COVID-19. Front Public Health. 2022;10: 985553.
50. Auf der Strasse W, Campos DP, Mendonça CJA, Soni JF, Mendes J, Nohama P. Forehead, temple and wrist temperature assessment of ethnic groups using infrared technology. Med Eng Phys. 2022;102:103777.
51. Norman DC, Yoshikawa TT. Fever in the elderly. Infect Dis Clin North Am. 1996;10(1):93–9.
52. McConeghy KW, White E, Panagiotou OA, Santostefano C, Halladay C, Feifer RA, Blackman C, Rudolph JL, Mor V, Gravenstein S. Temperature Screening for SARS-CoV-2 in Nursing Homes: Evidence from Two National Cohorts. J Am Geriatr Soc. 2020;68(12):2716–20.
53. Dell'Isola GB, Cosentini E, Canale L, Ficco G, Dell'Isola M. Noncontact body temperature measurement: uncertainty evaluation and screening decision rule to prevent the spread of COVID-19. Sensors (Basel). 2021;21(2):346.