Journal of Pediatric Nursing 46 (2019) 89–99

Contents lists available at ScienceDirect

Journal of Pediatric Nursing

journal homepage: www.pediatricnursing.org

Temporal Artery Thermometry in Pediatric Patients: Systematic Review and Meta-Analysis

Panagiotis Kiekkas, PhD, RN [a,*], Diamanto Aretha, PhD, MD [b], Eleni Almpani, PhD, RN [a], Nikolaos Stefanopoulos, PhD, RN [a]

[a] Nursing Department, Technological Educational Institute of Western Greece, Patras, Greece
[b] Department of Anesthesiology and Critical Care Medicine, Patras University Hospital, Patras, Greece

ARTICLE INFO

Article history:
Received 18 November 2018
Revised 4 March 2019
Accepted 4 March 2019

Keywords:
Temporal artery
Temperature measurement
Pediatric patients
Accuracy
Sensitivity
Specificity

ABSTRACT

Problem: Non-invasive thermometry methods have been used as substitutes for intra-corporeal ones in order to decrease patient discomfort and risk for complications, yet the evaluation of their performance is necessary. Our aim was to synthesize the evidence on the accuracy and precision of temporal artery (TA) thermometry, as well as on its sensitivity and specificity for fever detection.
Eligibility criteria: This systematic review and meta-analysis included method-comparison studies, which compared TA temperature measurements with invasive thermometry ones, were published between 2000 and 2018, and were conducted on patients aged <18 years.
Sample: Thirty articles were selected for inclusion in the final analysis after screening those identified by searches in CINAHL, PubMed, Web of Science, Cochrane Library, EMBASE and Scopus.
Results: Quantitative synthesis indicated that pooled mean TA temperature was lower than core temperature by 0.01 °C (95% limits of agreement, −0.06 °C to 0.03 °C). Average summary sensitivity and specificity for fever detection were 0.72 (95% confidence interval, 0.66–0.79) and 0.91 (95% confidence interval, 0.86–0.93) respectively. Subgroup analysis indicated a trend toward larger temperature underestimation in febrile patients and in ages ≤4 years.
Conclusions: Despite its satisfactory accuracy, precision and specificity, TA thermometry has low sensitivity when used in pediatric patients, which does not allow satisfactory fever detection.
Implications: TA thermometry cannot be recommended for replacing rectal temperature measurement methods in children, due to its high proportion of false negative readings during screening for fever.

Background

Core temperature (T_c) measurement has traditionally been an essential component of pediatric patients' healthcare, considering that fever detection is often indicative of various disease processes and guides diagnostic investigation and therapeutic decisions (Allegaert, Casteels, van Gorp, & Bogaert, 2014; Al-Mukhaizeem et al., 2004). Thermistor probes inserted in the pulmonary artery, urinary bladder, esophagus, nasopharynx and rectum have been used for providing valid T_c readings, by measuring blood temperature of vessels that perfuse head and trunk organs (Batra & Goyal, 2013). Among these invasive methods, rectal thermometry has been the gold-standard reference and the most frequently used method in infants and children (Allegaert et al., 2014). Unfortunately, intra-corporeal thermometry methods are generally characterized by difficulty and limitations in their use, relatively high cost and high risk for complications. With regard to probe or thermometer insertion in the rectum, reported disadvantages include discomfort, risk for traumatic rectum injury and transmission of stool-borne pathogens for patients, emotional upset for patients and families, while nurses consider rectal thermometry to be time-consuming and lead to treatment delays (Greenes & Fleisher, 2001; Holzhauer, Reith, Sawin, & Yen, 2009).

According to a Society of Pediatric Nurses' position statement for body temperature measurement "the goal is to use the most accurate method with the least degree of variance while still recognizing the comfort of the patient and ease of use for the healthcare provider" (Asher & Northington, 2008). In this context, non-invasive thermometry methods have been introduced to clinical practice for combining patient safety and comfort with satisfactory accuracy of T_c measurements. Non-invasive thermometry mainly includes oral, tympanic, temporal artery (TA) and axillary temperature measurement by digital electronic thermometers, which display estimated T_c values according to conversion algorithms (Jefferies, Weatherall, Young, & Beasley, 2011; Lawson et al., 2007). For this reason, accuracy and precision of non-invasive

* Corresponding author at:
E-mail address: kiekkpan@otenet.gr (P. Kiekkas).

https://doi.org/10.1016/j.pedn.2019.03.004
0882-5963/

methods are inferior to those of intra-corporeal ones (O'Grady et al., 2008). Accuracy refers to the degree of agreement between the readings of a new method and those of the standard method, while precision refers to the amount of variability between the values of compared methods (Barnason et al., 2012).

Considering that elevated T_c due to the febrile response is a common, early sign of many infectious and non-infectious pathologic conditions, failure to detect fever is expected to delay the diagnostic and treatment process (Moore, Carrigan, Solomon, & Tart, 2015). On the other hand, over-diagnosis of fever is followed by unnecessary diagnostic procedures or drug administration, cancellation of elective operations, and prolonged hospitalization (Mangat, Standley, Prevost, Vasconcelos, & White, 2010; Nimah, Bshesh, Callahan, & Jacobs, 2006). Therefore, besides accuracy and precision, primary criteria for establishing the use of non-invasive thermometry methods should be their satisfactory sensitivity and specificity for fever detection, so that their measurements do not misclassify fever (Hamilton, Marcos, & Secic, 2013). Sensitivity refers to the proportion of people with disease who will have a positive result, or to the proportion of febrile cases that are correctly classified as such by the thermometry method. Specificity refers to the proportion of people without the disease who will have a negative result, or to the proportion of non-febrile cases that are correctly classified as such by the thermometry method (Akobeng, 2006).

Among non-invasive methods, TA thermometry is based on multiple readings of the heat emitted as infrared radiation over the forehead and temporal region, in order to capture the highest value of TA blood temperature (Kimberger, Cohen, Illievich, & Lenhardt, 2007). T_c is then calculated by a proprietary algorithm, which compensates for ambient temperature (Marable, Shaffer, Dizon, & Opalek, 2009). Existing studies have confirmed that TA thermometry is overwhelmingly preferred by pediatric patients and their families thanks to its non-stressful nature (Opersteny et al., 2017), while the majority of nursing personnel trust its accuracy (Hurwitz, Brown, & Altmiller, 2015).

The objective of this review and meta-analysis was to determine whether TA thermometry has acceptable accuracy and precision for measuring T_c, as well as acceptable sensitivity and specificity for fever detection, in pediatric patients. Recently published systematic reviews and meta-analyses on the diagnostic accuracy of TA thermometry have included studies conducted on both adult and pediatric populations (Geijer, Udumyan, Lohse, & Nilsagard, 2016; Niven et al., 2015). However, considering the reported differences in the performance of TA thermometry between adults and children (Suleman, Doufas, Akça, Ducharme, & Sessler, 2002), the need for an updated systematic review and meta-analysis focused on pediatric patients was identified.

Aims

The aims of this systematic literature review and meta-analysis were to systematically search, identify, critically appraise and synthesize qualitatively and quantitatively the existing empirical evidence from studies conducted on pediatric patients (a) on the accuracy and precision of TA temperature measurements compared to T_c ones estimated by the use of invasive thermometry (reference standards), and (b) on sensitivity and specificity of TA temperature measurements for detecting fever.

Methods

This systematic review and meta-analysis was conducted following the guidelines set out in the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement (Liberati et al., 2009).

Inclusion and exclusion criteria

Articles published between January 2000 and October 2018 in English-language, peer-reviewed journals were considered for inclusion. All the following inclusion criteria should be met for selecting a study:

(a) Study subjects: pediatric patients (aged <18 years) admitted to any hospital department or unit.
(b) Study design: cross-sectional, prospective or retrospective, single- or multi-center, method-comparison studies, in which TA temperature measurements were compared with those of invasive thermometry, including pulmonary artery, urinary bladder, esophagus, nasopharynx or rectum temperatures.
(c) Publication types: original full articles.
(d) Statistical methods: (1) accuracy and precision of TA temperature measurements, assessed by comparisons based on the analysis recommended by Bland and Altman (1999), were reported, or reported data allowed their calculation, or alternatively, (2) sensitivity and specificity of TA temperature measurements for fever detection were reported, or reported data allowed their calculation.

Exclusion criteria were:

(a) Study subjects: adults only.
(b) Study design: interventional, non-clinical, or TA temperature measurements were compared only with those of non-invasive thermometry (oral, tympanic or axillary).
(c) Publication types: letters or congress abstracts.

Database search and study selection

The third and fourth author independently and systematically searched for studies indexed in the Cumulative Index for Nursing and Allied Health Literature (CINAHL), the US National Library of Medicine (PubMed), the Web of Science, the Cochrane Library, the Excerpta Medica Database (EMBASE), and Scopus. The search terms “temporal artery”, “forehead”, “thermometry”, “thermometer”, “core temperature”, “temperature measurement” were generally combined with “pediatric”, “children”, “neonatal” and “infant”. Database searches took place at the first week of October 2018.

Study selection was conducted in three steps. In the first step, articles were screened for inclusion according to their titles and abstracts. In the second step, full text of selected articles was obtained and read for determining eligibility for inclusion. In the third step, reference lists of eligible articles were checked for potentially relevant articles not found in the online searches, and their full text was also read for determining eligibility for inclusion. Discrepancies between reviewers were discussed until consensus was reached.

Data extraction and quality appraisal

In included studies, temperature values were graphically represented on a Bland-Altman plot (as difference between methods against their mean). Accuracy corresponded to the mean difference (systematic error or bias) between the values of TA and reference standard temperature measurements; bias quantified how much higher (positive bias) or lower (negative bias) TA temperature values were compared with reference standard ones. Precision corresponded to the 95% limits of agreement (LoA) (random error, mean difference ± 1.96 standard deviation) between these values (Bland & Altman, 1999; Hanneman, 2008).

The first and second author independently extracted data from included studies by using a standardized data collection form. Extracted data included:

(a) study description: population characteristics (age range, gender, department or unit of admission), temperature measurement

sites, types of thermometers used, T_c range (measured by reference standards), percentage of febrile patients,

(b) study outcomes: mean difference and 95% LoA between TA and reference standard temperature measurements, sensitivity and specificity of TA temperature measurements for fever detection. In case accuracy, precision, sensitivity and/or specificity were not directly reported, these were computed from reported data according to standard formulae.

Furthermore, the first and second author evaluated methodological quality of included studies by using QUADAS-2 (Whiting et al., 2011), which has been developed for evaluating quality of method-comparison studies. In addition, suggested criteria for quality appraisal of method-comparison studies, especially of those comparing thermometry methods (Bland & Altman, 1999; Hanneman, 2008; Hooper & Andrews, 2006; Kottner et al., 2011), were adopted for further evaluation. These criteria included: (a) conduction of power analysis, (b) sample size and number of paired temperature measurements included in the Bland-Altman analysis, (c) repeated temperature measurements on the same patient included in the Bland-Altman analysis, (d) data collector information, including training before TA thermometer use, number of collectors and interrater reliability in case of multiple collectors, and (e) time between paired temperature measurements.

Collected data on study description, outcomes and quality appraisal were compared, and in cases where there was discordance between reviewers, the articles were re-examined until differences were resolved by consensus. Study description and quality appraisal criteria were summarized in a table and within the text.

Data analysis

Quantitative synthesis of data from included studies was conducted with Comprehensive Meta Analysis 3.3 software (Biostat, Englewood, NJ). For each individual comparison, mean difference and 95% LoA between TA and T_c temperature measurements, as well as sensitivity and specificity of TA thermometry for fever detection with their 95% confidence intervals (CIs), were used as outcome measures.

A common effect size could not be assumed for included studies, since these have been conducted on diverse patient populations and clinical settings, different thermometers and reference standards have been used, and there has been wide variation in the way temperature measurements were conducted. Thus, a random effects approach was used for combining individual study estimates to obtain pooled estimates of systematic error (bias - mean difference) and random error (95% LoA) (Tipton & Shuster, 2017; Williamson, Lancaster, Craig, & Smyth, 2002). Likewise, average summary estimates of sensitivity and specificity for fever detection were calculated by using a

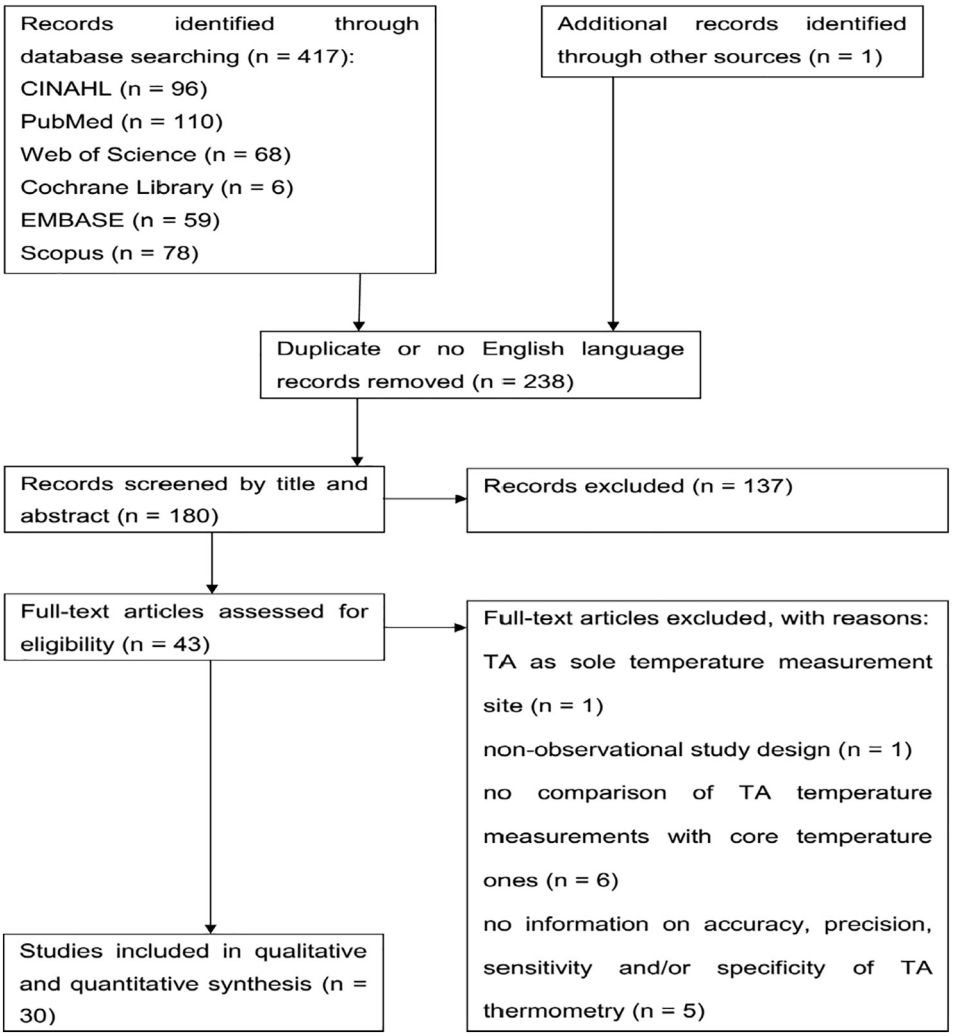

Fig. 1. Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) flow diagram (TA = temporal artery).

Table 1
Description of included studies.

Study: author(s) (year), country	Study population characteristics	Power analysis/number of paired TMs per pt (total)[a]	T_c range/% febrile[b]	Standard TM site (brand)/TA thermometer	Time between paired TMs	Training before TMs/data collectors/interrater reliability
Allegaert et al. (2014), Belgium	294 pediatric ward pts, age ≤17y, median age 3.2y	Not reported/1 (294)	35.1–39.5 °C/7.5% (T_c > 38 °C)	Rectal (Filac 3000, Covidien)/TAT-5000	<5 min	Yes/ward nurses/not reported
Al-Mukhaizeem et al. (2004), Canada	80 elective dental surgery pts, mean age 45 ± 35 m	Yes/1 (80)	36.6 ± 0 .4 °C/2.5% (T_c > 38 °C)	Oesophageal (TeleThermometer, YSI)/LXTA Temporal scanner	Not reported	Yes/one nurse/–
Bahorski et al. (2012), USA	47 emergency, ICU and outpatient unit pts, age range 3–36 m, 43% male	Yes/1 (47)	36.1–40.2 °C/46.8% (T_c ≥ 38 °C)	Rectal (SureTemp, WelchAllyn)/TAT-5000	Rapid sequential	Yes/3 nurses/not reported
Batra and Goyal (2013), India	100 emergency pts, age range 2–12y, mean age 6.1y, 54% male	Yes/1 (100)	37.3 ± 0 .4 °C in non-febrile, 39.4 ± 0 .9 °C in febrile/50% (T_c > 38 °C)	Rectal (Hicks Thermometers)/TAT-2000C	Not reported	Not reported/not reported/–
Callanan (2003), USA	187 emergency pts, age <3 m	Not reported/not reported	Not reported/12.3% (T_c > 38 °C)	Rectal (SureTemp, WelchAllyn)/SensorTouch TA	Not reported	Not reported/not reported/–
Carr et al. (2011), USA	40 hospital pts, age ≤24 m, mean age 10.9m, 55% male	Not reported/5–18 (450)	Not reported/100% (T_c > 38 °C)	Rectal (SureTemp, WelchAllyn)/TAT-5000	Not reported	Yes/all nursing personnel/r = 0.98
Drake-Brockman, Hegarty, Chambers, and von Ungern-Sternberg (2014), Australia	200 elective non-cardiac surgery pts, age ≤17y, mean age 8.4 ± 0.2y, 59% male	Not reported/not reported	36.3 ± 0 .6 °C/0%	Nasopharyngeal (IntelliVue MP800, Philips)/TAT-5000	Simultaneous	Not reported/not reported/–
Evron et al. (2017), Finland	16 elective surgery pts., age <5y, median age 0.8 (0.4–1.6)y	Not reported/median 51	35.2–37.8 °C/0%	Oesophageal (Novamed)/Temple Touch Pro sensor (Medisim)	Simultaneous	Not reported/not reported/–
Forrest et al. (2017), USA	85 emergency pts, age <36 m, mean age12 ± 8 m, 49% male	Not reported/1 (85)	Mean 37.6 °C/30.6% (T_c > 38 °C)	Rectal (SureTemp, WelchAllyn)/TAT-5000	<5 min	Yes/6 investigators/not reported
Greenes and Fleisher (2001), USA	304 emergency pts, age <1y	Not reported/1 (304)	35.7–40.7 °C/35.9% (T_c > 38 °C)	Rectal (Diatek, Welch Allyn)/LXTA Temporal scanner	Not reported	Yes/research assistants/not reported
Gunawan et al. (2010), Indonesia	134 neonatal pts, age >24 h, mean age 36 ± 13 h, 52% male	Yes/1 (134)	36.4–37.8 °C/0%	Rectal (Clinical thermometer-CE 0197)/TAT-5000	Not reported	Yes/1 medical doctor/–
Hamilton et al. (2013), Argentina	205 emergency, pediatric ICU and other pts, age ≤18y, mean age 66 m, 58% male	Not reported/1 (205)	Not reported/45.9% (T_c > 38 °C)	Rectal (SureTemp Plus, WelchAllyn)/TAT-5000	Rapid sequential	Yes/one nurse/–
Hebbar, Fortenberry, Rogers, Merritt, and Easley (2005), USA	44 pediatric ICU pts, mean age 11.5m (interquartile range 2–34 m)	Yes/1 (44)	37.6 ± 1 .1 °C/43.2% (T_c > 38 °C)	Rectal (Allegiance Healthcare Corp.)/LXTA Temporal scanner	Rapid sequential	Yes/all nurses/not reported
Hoffman et al. (2013), USA	263 emergency pts, age <36 m, 48% male	Not reported/1 (263)	Mean 39.1 °C/55.9% (T_c ≥ 38 °C)	Rectal (Alaris Medical Sciences)/TAT-2000C	Mean 17.2min	Not reported/one investigator/–

Holzhauer et al. (2009), USA	474 emergency pts, age range 3–36 m	Yes/1 (474)	37.7 ± 0 .9 °C/42.4% (T_c > 38 °C)	Rectal (WelchAllyn)/TAT-5000	Not reported	Yes/all nurses/not reported
Lee et al. (2011), USA	34 neonatal ICU pts, mean age 35.7 ± 1.8w, 53% male	Yes/1 (34)	35.9–37.2 °C/0%	Rectal (Smiths Medical ASD)/TAT-5000	<2 min	Yes/8 investigators/not reported
Mogensen, Wittenhoff, Fruerhoj, and Hansen (2018), Denmark	965 emergency pts, age <18y, median age 24 m (interquartile range 11–70 m), 51% male	Yes/1 (965)	Not reported/39.0% (T_c > 38 °C)	Rectal (OMRON Healthcare)/TAT-5000	Simultaneous	Yes/all nurses/not reported
Moore et al. (2015), USA	239 emergency pts, age range 3 m–4y, mean age 1.5 ± 0.8y, 53% male	Yes/1 (239)	38.1 ± 1 .0°C/41.0% (T_c ≥ 38 °C)	Rectal (Alaris Medical Sciences)/TAT-5000	Rapid sequential	Yes/5 nurses/not reported
Nimah et al. (2006), USA	36 pediatric and cardiac ICU pts, age ≤77 m, mean age 20 ± 18.6m, 58% male	Not reported/multiple (902)	Not reported/47.2% (T_c ≥ 38 °C)	Urinary bladder (400 Series thermistor)/SensorTouch TM	Rapid sequential	Yes/2 nurses/not reported
Odinaka, Edelu, Nwolisa, Amamilo, and Okolo (2014), Nigeria	156 emergency pts, age ≤59 m, mean age 10.8 ± 13.6m, 52% male	Not reported/1 (156)	35.7–40.5 °C/50.6% (T_c ≥ 38 °C)	Rectal (mercury in glass)/TAT-2000C	Simultaneous	Not reported/not reported/–
Paes, Vermeulen, Brohet, van der Ploeg, and de Winter (2010), The Netherlands	100 pediatric ward pts, age <18y, mean age 3.2y, 50% male	Not reported/1 (100)	35.9–40.2 °C/50.0% (T_c ≥ 38 °C)	Rectal (Terumo Corp.)/Thermofocus	<5 min	Not reported/2 investigators/not reported
Penning, van der Linden, Tibboel, and Evenhuis (2011), The Netherlands	198 emergency and day-care pts, age ≤18y, mean age 5.1 ± 4.7y, 61% male	Not reported/1 (198)	35.9–38.2 °C (day-care), 37.7–40.4 °C (emergency)/40.9% (T_c > 38 °C)	Rectal (Terumo Corp.)/TAT-5000	<15 min	Yes/all nurses/not reported
Reynolds et al. (2014), USA	52 emergency pts, age <4y, mean age 13.5 ± 8 .0m, 60% male	Yes/1 (52)	36.6–40.1 °C/34.6% (T_c ≥ 38 °C)	Rectal (SureTemp Plus, WelchAllyn)/TAT-5000	Not reported	Yes/11 investigators/not reported
Sahin et al. (2012), Turkey	60 elective lower abdominal surgery pts, age range 1 m–4y, 45% male	Not reported/7 (420)	Not reported/0%	Nasopharyngeal (GE Datex-Ohmeda)/PlusMRD TA Thermometer	<5 min	Not reported/anesthesiologists/not reported
Schuh et al. (2004), Canada	327 emergency pts, age <24 m, mean age 9.2 ± 6 .8m, 45% male	Yes/1 (327)	Not reported/41.6% (T_c ≥ 38 °C)	Rectal (Alaris Medical Sciences)/LXTA Temporal scanner	Not reported	Yes/one nurse/–
Selent et al. (2013), USA	855 emergency pts, age range 6 m–17y, 55% male	Not reported/not reported	38.5 (36.2–41.1)/46.8% (T_c ≥ 38 °C)	Rectal (Alaris Medical Sciences)/Thermofocus	Not reported	Not reported
Siberry, Diener-West, Schappell, and Karron (2002), USA	275 acute care pts, age <24 m, mean age 11.2m, 49% male	Not reported/1 (275)	35.6–40.9 °C/17.1% (T_c ≥ 38 °C)	Rectal (SureTemp Plus, WelchAllyn)/LXTA Temporal scanner	Not reported	Yes/5 nurses/not reported
Suleman et al. (2002), USA	26 surgical pts recovering from CPB, mean age 3.0 ± 4.0y, 60% male	Not reported/multiple (246)	38.0 ± 0 .8 °C/62.5% (T_c ≥ 37.8 °C)	Urinary bladder/SensorTouch TM	Simultaneous	Yes/1 investigator/–
Teran et al. (2012), Bolivia	434 emergency and inpatient unit pts, age <48 m, mean age 14.6 ± 10.7m, 48% male	Not reported/1 (434)	37.9 ± 0 .9 °C/38.5% (T_c ≥ 38 °C)	Rectal (mercury in glass)/TAT-2000C	≤15 s	Not reported/not reported/–
Titus, Hulsey, Heckman, and Losek (2009), USA	42 emergency pts, age range 1–4y	Not reported/not reported	36.0–40.6 °C/26.2% (T_c ≥ 38.3 °C)	Rectal (Turbo Temp, Alaris)/TAT-5000	Not reported	Yes/4 research assistants/not reported

TM = temperature measurement; TA = temporal artery; ICU = intensive care unit; CPB = cardiopulmonary bypass; T_c = core temperature; pts = patients; y = years; m = months; w = weeks; d = days; min = minutes; s = seconds.

[a] Measurements between temporal artery and reference standard, included in the Bland-Altman analysis.

[b] According to reference standard.

bivariate, random effects approach, which jointly analysed pairs of sensitivity and specificity to account for the patterns of correlation between them (Reitsma et al., 2005). Forest plots were constructed to display individual and pooled estimates of mean difference and 95% LoA, as well as individual and average summary estimates of sensitivity and specificity.

Heterogeneity of mean differences, standard deviations, sensitivity and specificity across studies was evaluated by calculating the Cochran Q test and the I^2 statistic; substantial heterogeneity was indicated by $p < 0.05$ and I^2 value >50%. Subgroup analyses were performed according to patient age, febrile status, reference standard and TA thermometer, including studies that reported relevant data. Publication bias was not assessed, since respective standard tests are not recommended for meta-analysis of diagnostic accuracy studies (Geijer et al., 2016).

Results

Search outcome

Searches revealed 418 potentially relevant citations (Fig. 1). Removal of duplicate and non-English language citations, along with screening of titles and abstracts, yielded 43 articles for full-text review. Of them, 30 articles (conducted on 30 unique study populations) met eligibility criteria for inclusion in the qualitative and quantitative synthesis.

Study description and quality appraisal

Characteristics of included studies are presented in Table 1. All studies were published between 2001 and 2018, were single-center and prospective, with the exception of one study with retrospective design (Hoffman, Etwaru, Dreisinger, Khokhar, & Husk, 2013). Studies mainly enrolled emergency department (n = 17), surgical (n = 5), and intensive care unit (n = 4) patients. Thermometers used for measuring TA temperature included TAT-5000 (Exergen Corp., Watertown, MA) (n = 14), LXTA Temporal scanner (Exergen Corp., Watertown, MA) (n = 5), TAT-2000C (Exergen Corp., Watertown, MA) (n = 4), SensorTouch TM (Philips Electronic Corp., Chicago, IL) (n = 2), Thermofocus (Technimed, Varese, Italy) (n = 2), SensorTouch TA (Exergen Corp., Watertown, MA) (n = 1), Temple Touch Pro (Medisim, Neve Ilan, Israel) (n = 1), and PlusMED TA (PlusMED, İstanbul, Turkey) (n = 1). Reference standards included rectal (n = 24), nasopharyngeal (n = 2), oesophageal (n = 2), and urinary bladder (n = 2) temperature. The majority of studies (n = 24) enrolled both febrile and non-febrile patients.

Evaluation of included studies according to quality appraisal criteria is presented in Table 1. Only few studies (n = 11) conducted power analysis for determining required sample size. Most studies (n = 14) enrolled relatively small sample sizes of ≤100 patients; few studies (n = 5) enrolled 101–200 patients, while the rest (n = 11) enrolled >200 patients. In most studies (n = 21), only one paired temperature measurement (between TA and reference standard temperature) per patient was included in Bland-Altman analysis. In few studies (n = 5), repeated measurements on the same patient were included in Bland-Altman analysis; this allowed a larger number of temperature measurements, but may have resulted in the underestimation of measurement variation. In most studies, data collectors (physicians, nurses or investigators) were reported to be trained in the appropriate use of TA thermometer (n = 20), and be more than one (n = 17); however, interrater reliability was rarely evaluated (n = 1). In some studies, temperature measurements were simultaneous (n = 5) or rapid sequential (n = 5), while in others, time between measurements was <5 min (n = 4), <2 min (n = 1), and ≤15 s (n = 1); in the rest studies (n = 14), time between measurements was either not reported or >5 min.

QUADAS-2 assessments are presented in Fig. 2. In summary, risk of bias was assessed as low for patient selection (n = 8), index test (n = 29), reference standard (n = 28), and flow and timing (n = 12) domain. Applicability concerns were assessed as low for patient selection, index test and reference standard domain (n = 30).

Quantitative synthesis of study findings

Data on the accuracy and precision of TA temperature measurements were extracted by 27 studies, which enrolled 5178 pediatric patients in total. Reference standards included rectal (n = 21), nasopharyngeal (n = 2), oesophageal (n = 2), and urinary bladder (n = 2) temperature. Mean differences and 95% LoA of temperature readings are displayed in Fig. 3, with some studies (n = 2) providing estimates for patient subgroups. In the majority of studies (n = 21), TA temperature measurements were lower than T_c ones. The overall pooled mean difference between TA temperature and T_c was −0.01 °C (95% LoA, −0.06 °C to 0.03 °C). Inter-study heterogeneity was not substantial ($Q = 51.5$, $p = 0.004$, $I^2 = 45.6\%$). Subgroup analyses are presented in Table 2. Trends for larger differences from the reference (indicating larger T_c underestimation) were identified for febrile patients and those aged ≤4 years. Inter-study heterogeneity was substantial for rectal temperature ($Q = 48.5$, $p < 0.001$, $I^2 = 58.8\%$) and TAT-5000 ($Q = 43.7$, $p < 0.001$, $I^2 = 70.2\%$).

Data on sensitivity and specificity of TA temperature measurements for fever detection were extracted by 21 studies, which enrolled 5619 pediatric patients in total (2163 of them manifested fever and 3456 were non-febrile). Temperature thresholds used for fever definition included $T_c \geq 38$ °C (n = 19), $T_c \geq 37.8$ °C (n = 1), and $T_c \geq 38.3$ °C (n = 1). Reference standards included rectal (n = 19) and urinary bladder (n = 2) temperature. Estimates and 95% CIs for sensitivity and specificity of TA temperature measurements for fever detection are displayed in Fig. 4. Sensitivity estimates varied between 0.41 and 1.00, with average summary sensitivity being 0.72 (95% CI, 0.66–0.79). Specificity estimates varied between 0.46 and 1.00, with average summary specificity being 0.91 (95% CI, 0.86–0.93). Average summary diagnostic odds ratio, positive likelihood ratio and negative likelihood ratio were 25.81 (95% CI, 20.48–28.23), 8.00 (95% CI, 4.71–11.29), and 0.31 (95% CI, 0.23–0.40) respectively. Inter-study heterogeneity was substantial for both sensitivity ($Q = 549.6$, $p < 0.001$, $I^2 = 96.4\%$) and specificity ($Q = 346.9$, $p < 0.001$, $I^2 = 94.2\%$). Average summary estimates for sensitivity and specificity did not differ considerably after excluding studies (n = 2) that did not use $T_c \geq 38$ °C as fever definition threshold, being 0.71 (95% CI, 0.65–0.76) and 0.91 (95% CI, 0.87–0.93) respectively.

Discussion

The findings of this meta-analysis showed that TA thermometry in pediatric patients underestimated T_c by a pooled mean difference of −0.01 °C, with 95% LoA of only ±0.05 °C (−0.06 to 0.03 °C). An accuracy standard of ±0 .5 °C has been most commonly used as clinically acceptable deviation from the reference temperature (Barnason et al., 2012; Gunawan, Soetjiningsih, & Kardana, 2010; Kimberger et al., 2007). Apparently, 95% LoA of TA thermometry were much narrower compared to this accuracy standard, for both overall pediatric population and patient subgroups, being ±0.17 °C for febrile patients, ±0.06 °C for children ≤2 years, and ±0.04 °C for children ≤4 years. These findings indicate that accuracy and precision of TA thermometry are considered quite satisfactory and do not preclude its use in pediatric patients.

Accuracy and precision of TA thermometry were considerably higher compared to those of other non-invasive thermometry methods in children. Previous meta-analyses have reported mean differences of −0.22 °C (95% LoA, −0.44 to 1.30 °C) between tympanic temperature and rectal one (Zhen, Xia, Long, & Pu, 2014), as well as of −0.17 °C (95% LoA, −0.15 to 0.50 °C) between axillary temperature and rectal one (Craig, Lancaster, Williamson, & Smyth, 2000). In addition, accuracy

Study	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Allegaert (2014)	+	+	+	+	+	+	+
Al-Mukhaizeem (2004)	?	+	+	?	+	+	+
Bahorski (2012)	?	+	+	+	+	+	+
Batra (2013)	?	+	+	?	+	+	+
Callanan (2003)	?	+	+	?	+	+	+
Carr (2011)	?	+	+	?	+	+	+
Drake-Brockman (2014)	?	+	+	+	+	+	+
Evron (2017)	?	+	+	?	+	+	+
Forrest (2017)	+	+	+	+	+	+	+
Greenes (2001)	?	+	+	?	+	+	+
Gunawan (2010)	?	+	+	?	+	+	+
Hamilton (2013)	?	+	+	+	+	+	+
Hebbar (2005)	+	?	?	+	+	+	+
Hoffman (2013)	+	+	+	?	+	+	+
Holzhauer (2009)	?	+	+	?	+	+	+
Lee (2011)	?	+	+	+	+	+	+
Mogensen (2018)	+	+	+	?	+	+	+
Moore (2015)	?	+	+	?	+	+	+
Nimah (2006)	+	+	+	+	+	+	+
Odinaka (2014)	?	+	-	+	+	+	+
Paes (2010)	?	+	+	+	+	+	+
Penning (2011)	?	+	+	+	+	+	+
Reynolds (2014)	?	+	+	?	+	+	+
Sahin (2012)	+	+	+	+	+	+	+
Schuh (2004)	?	+	+	?	+	+	+
Selent (2013)	?	+	+	?	+	+	+
Siberry (2002)	?	+	+	?	+	+	+
Suleman (2002)	-	+	+	?	+	+	+
Teran (2012)	+	+	+	?	+	+	+
Titus (2009)	-	+	+	?	+	+	+

Fig. 2. Quality assessment of included studies with QUADAS-2: summary for risk of bias and applicability concerns (− = high risk/concerns; + = low risk/concerns; ? = unclear).

and precision of TA thermometry in children were considerably higher compared to those reported for adults, which corresponded to a mean difference of −0.17 °C (95% LoA, −1.14 to 0.79 °C) (Geijer et al., 2016). Possible explanations might be the thinner skin layer over superficial TA and frontal bone in childhood, along with the absence of atherosclerotic TA changes (Bodkin, Acquisto, Zwart, & Toussaint, 2014; Sahin et al., 2012).

However, accuracy and precision of TA thermometry calculated in the present quantitative synthesis contradict the findings of a meta-analysis, which reported a considerably larger mean difference of −0.20 °C (95% LoA, −1.17 to 0.76 °C) between TA temperature measurements and T_c ones obtained by invasive thermometry in children (Geijer et al., 2016). This discrepancy can be a result of differences in included studies. Two studies included in the previous quantitative synthesis were not included in the present one (the first due to non-English language and the second due to inappropriate design). At the same time, the present quantitative synthesis included five studies not included in the previous one, which enrolled 1429 patients (about 28% of total).

TA temperature measurements had high pooled specificity but low sensitivity for detecting fever. A meta-analysis on TA thermometry for both adults and children reported similar findings of 0.72 (95% CI, 0.61–0.81) for sensitivity and 0.94 (95% CI, 0.87–0.97) for specificity (Geijer et al., 2016). Likewise, pooled sensitivity of 0.64 (95% CI, 0.55–0.72) and specificity of 0.96 (95% CI, 0.93–0.97) have been estimated for peripheral temperature measurements (including TA ones) in adults and children (Niven et al., 2015). In this case, TA thermometry is expected to result in a higher proportion of false-negative than false-positive readings. However, screening for fever constitutes the most important reason for measuring T_c in the pediatric population (Holzhauer

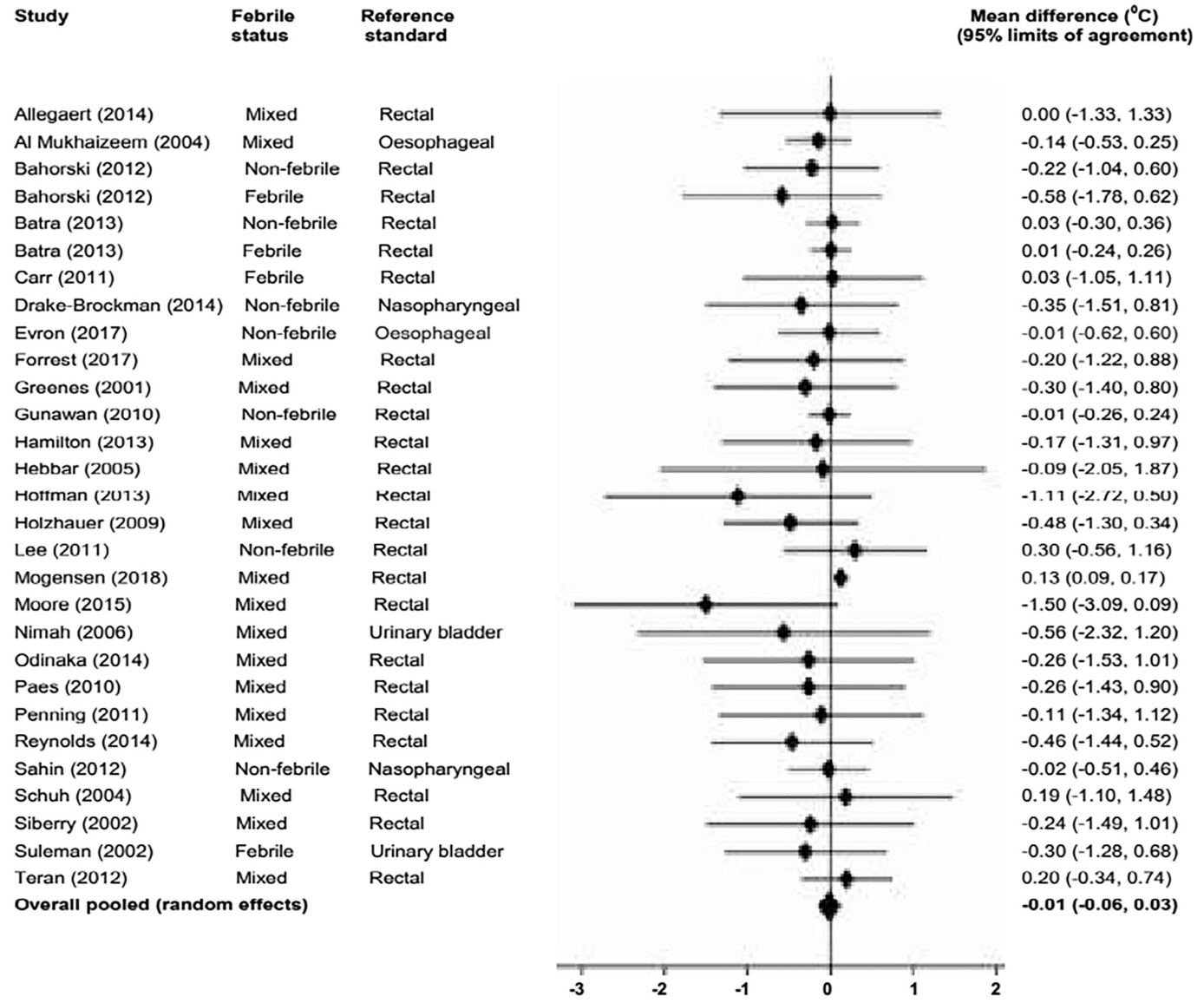

Fig. 3. Forest plot depicting individual and pooled mean temperature differences and 95% limits of agreement (temporal artery temperature vs. core temperature). Positive/negative values indicate that temporal artery temperature measurements overestimated or underestimated respectively core temperature ones. Two studies (Bahorski and Batra) reported separate estimates of mean difference for febrile and non-febrile patient subgroups.

et al., 2009). In this context, false-negative readings are generally more concerning than false-positive ones, since they are expected to be followed by false reassurance and sense of safety, or even by potentially harmful delay in providing necessary treatment (Schuh et al., 2004). For this reason, most clinicians would not accept the use of thermometry methods that fail to detect >10% of fever cases (Holzhauer et al., 2009; Schuh et al., 2004).

Table 2
Subgroup analysis: pooled mean differences and 95% limits of agreement (temporal artery temperature vs. core temperature).

	Mean difference (°C)	95% limits of agreement (°C)
Febrile status		
Febrile patients (n = 5)	−0.08	−0.24, 0.09
Non-febrile patients (n = 7)	−0.01	−0.10, 0.07
Reference standard		
Rectal temperature (n = 21)	−0.01	−0.06, 0.04
Patient age		
≤4 years (n = 16)	−0.03	−0.07, 0.01
≤2 years (n = 7)	−0.01	−0.06, 0.05
Temporal artery thermometer		
TAT-5000 (n = 13)	−0.01	−0.09, 0.06

High accuracy and precision of TA thermometry were combined with low sensitivity for fever detection. This finding is possibly not surprising, since data used for estimating accuracy, precision and sensitivity were extracted from different studies; some studies (n = 9) that provided data on accuracy and precision did not provide data on sensitivity, while others (n = 3) that provided data on sensitivity did not provide data on accuracy and precision. In addition, reference standard temperature measurements were more underestimated by TA ones at febrile temperatures [with mean difference and 95% LoA being −0.08 °C (−0.24 °C to 0.09 °C) vs. −0.01 °C (−0.10 °C to 0.07 °C) for non-febrile ones]. As has been suggested, accuracy of non-invasive temperature measurements can be negatively affected by physiologic responses related to fever, such as shivering during its chill phase and diaphoresis during its defervescence phase (Lawson et al., 2007; Mangat et al., 2010). Underestimation of T_c measurements at febrile temperatures may favor false-negative readings, accounting thus for the low sensitivity of TA thermometry for fever detection.

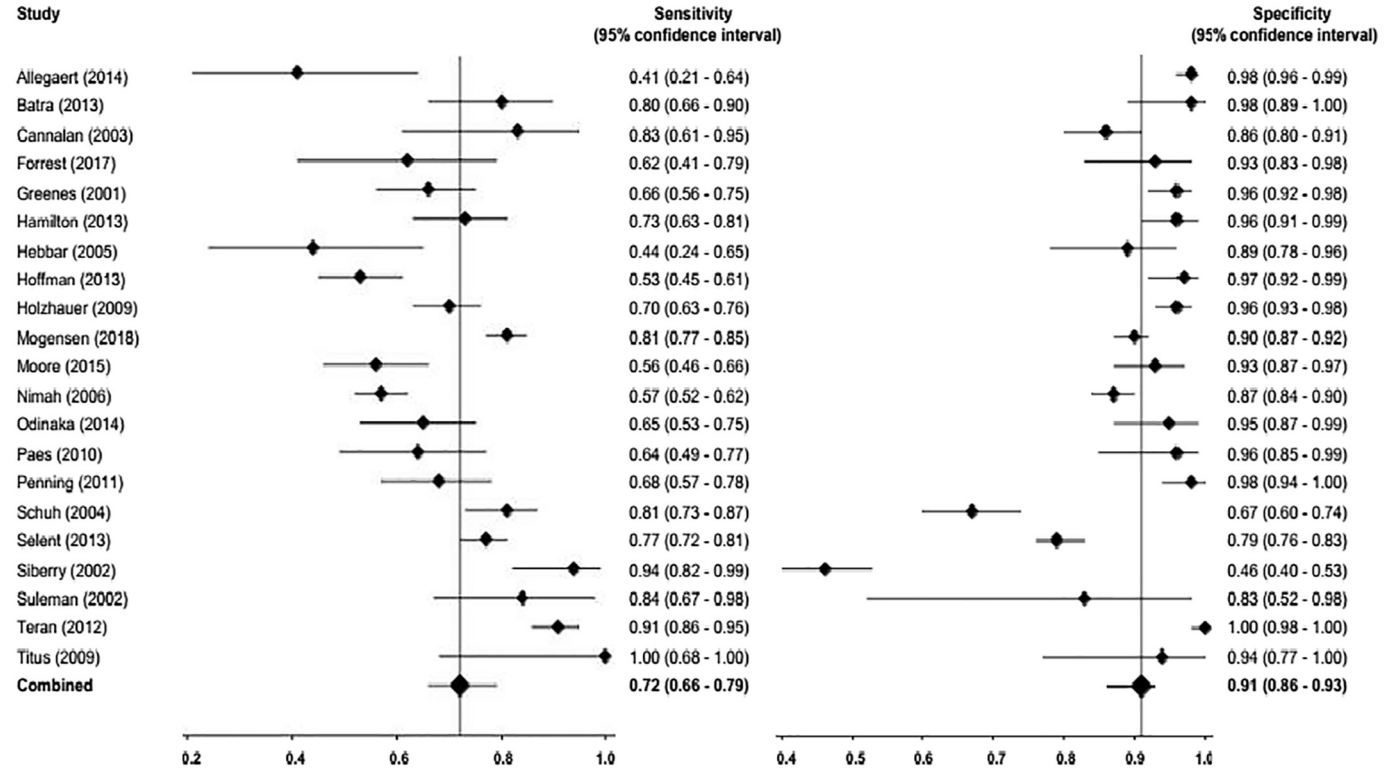

Fig. 4. Coupled forest plots depicting individual and pooled sensitivity and specificity (with 95% confidence intervals) of temporal artery temperature measurements for fever detection.

Rectal temperature constitutes the most frequently used thermometry method in children and was used as reference standard in the majority of included studies. However, the rectum is surrounded and insulated by large muscle mass, which may act as heat reservoir, thus its response to rapid T_c changes is expected to be slow (Al-Mukhaizeem et al., 2004; Carr et al., 2011). In febrile infants, TA temperature was found to decrease faster than rectal one after the administration of antipyretics (Greenes & Fleisher, 2004), which possibly means that TA temperature may better reflect current T_c than rectal one. In this case, rectal temperature could not be appropriate as reference standard for determining sensitivity and specificity of peripheral temperature measurements.

Study limitations

Most studies included in this systematic review and meta-analysis had methodological weaknesses, such as single-center design, small sample sizes, and inclusion of specific subgroups (e.g. neonatal patients); thus, generalizability of their findings may be limited. With regard to systematic review methodology, searches were conducted in a limited number of databases and only full-text, English-language articles were included; thus, other information sources were not covered. Moreover, search strategy was not determined and implemented in consultation with a librarian. With regard to quantitative synthesis of mean differences and 95% LoA, most subgroup analyses were based on limited data since the majority of studies did not provide separate values for subgroups. Quantitative synthesis of sensitivity and specificity could also be limited by the large heterogeneity among included studies. Finally, a summary receiver operating characteristics curve for sensitivity and specificity was not constructed.

Implications for clinical practice and research

In pediatric care, TA thermometry guarantees convenience and safety of use for children, ease and fastness of use for clinicians along with satisfactory accuracy and precision, while its high specificity ensures that fever can be safely confirmed in the majority of children found to be febrile by the use of TA thermometer. However, TA temperature measurements are currently characterized by unacceptably low sensitivity. This means that fever cannot be ruled out with satisfactory certainty in children found to be non-febrile by the use of TA thermometer, and a significant proportion of febrile episodes are expected to be missed when this site is used. Thus, TA thermometry cannot be considered appropriate as a substitute for invasive thermometry in patient groups with high incidence of fever, since the absolute number of false negative readings is expected to be high, or in those that appropriate fever treatment could significantly determine their outcomes. Furthermore, low sensitivity does not advocate for widely screening general pediatric population for fever by the use of TA thermometry. In case clinicians decide to use TA thermometry for fever screening, they should be aware of its high possibility for false negative readings, especially when febrile temperatures are approximated, e.g. between 37.5 and 37.9 °C; such temperature values are recommended to be further checked by invasive thermometry methods. It is also worth noticing that, in real clinical conditions, inappropriate personnel training, time constraints, and inadequate maintenance or calibration of devices can be followed by more inaccurate measurements of TA thermometry and even lower sensitivity than reported in studies.

Three main issues can be recommended for future research. First, considering that rectal temperature is possibly not the best choice for reflecting T_c, more data on the accuracy and precision, along with sensitivity and specificity for fever detection, of TA temperature measurements according to more valid reference standards, such as pulmonary artery, urinary bladder and esophagus, would be of importance. Second, low sensitivity of TA thermometry highlights the need for developing and evaluating the performance of improved devices, which could provide measurements of higher sensitivity, less affected by physiologic responses related to fever or ambient temperature. Third, only a small number of studies have currently provided data for specific patient subgroups (such as neonates or children <2 years),

which are characterized by thermoregulatory differences compared to older ages. Appropriateness of TA thermometry is therefore suggested to be separately studied for these patients.

Conclusions

The findings of this systematic review and meta-analysis revealed that TA thermometry has satisfactory accuracy, precision and specificity when used in pediatric patients. However, due to its low sensitivity, this method does not meet the demands of clinicians for identifying a satisfactory proportion of fever cases and cannot be recommended for the substitution of existing invasive methods of measuring T_c in children. Until improved devices are developed, clinicians should interpret TA thermometry readings with particular caution.

Declarations of interest

None.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

CRediT authorship contribution statement

Panagiotis Kiekkas: Conceptualization, Data curation, Formal analysis, Methodology, Supervision, Writing - original draft, Writing - review & editing. **Diamanto Aretha:** Data curation, Formal analysis, Writing - original draft, Writing - review & editing. **Eleni Almpani:** Investigation, Methodology, Validation, Visualization, Writing - review & editing. **Nikolaos Stefanopoulos:** Investigation, Methodology, Validation, Visualization, Writing - review & editing.

References

Akobeng, A. K. (2006). Understanding diagnostic tests 1: Sensitivity, specificity and predictive values. *Acta Paediatrica*, *96*, 338–341.

Allegaert, K., Casteels, K., van Gorp, I., & Bogaert, G. (2014). Tympanic, infrared skin, and temporal artery scan thermometers compared with rectal measurement in children: A real-life assessment. *Current Therapeutic Research*, *76*, 34–38.

Al-Mukhaizeem, F., Allen, U., Komar, L., Naser, B., Roy, L., Stephens, D., et al. (2004). Comparison of temporal artery, rectal and esophageal core temperatures in children: Results of a pilot study. *Paediatrics and Child Health*, *9*, 461–465.

Asher, C., & Northington, L. (2008). Position statement for measurement of temperature/fever in children. *Journal of Pediatric Nursing*, *23*, 234–236.

Bahorski, J., Repasky, T., Ranner, D., Fields, I., Jackson, M., Moultry, L., et al. (2012). Temperature measurement in pediatrics: A comparison of the rectal method versus the temporal artery method. *Journal of Pediatric Nursing*, *27*, 243–247.

Barnason, S., Williams, J., Proehl, J., Brim, C., Crowley, M., Leviner, S., et al. (2012). Emergency nursing resource: Non-invasive temperature measurement in the emergency department. *Journal of Emergency Nursing*, *38*, 523–530.

Batra, P., & Goyal, S. (2013). Comparison of rectal, axillary, tympanic, and temporal artery thermometry in the pediatric emergency room. *Pediatric Emergency Care*, *29*, 63–66.

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, *8*, 135–160.

Bodkin, R. P., Acquisto, N. M., Zwart, J. M., & Toussaint, S. P. (2014). Differences in noninvasive thermometer measurements in the adult emergency department. *American Journal of Emergency Medicine*, *32*, 987–989.

Callanan, D. (2003). Detecting fever in young infants: Reliability of perceived, pacifier, and temporal artery temperatures in infants younger than 3 months of age. *Pediatric Emergency Care*, *19*, 240–243.

Carr, E. A., Wilmoth, M. L., Eliades, A. B., Baker, P. J., Shelestak, D., Heisroth, K. L., et al. (2011). Comparison of temporal artery to rectal temperature measurements in children up to 24 months. *Journal of Pediatric Nursing*, *26*, 179–185.

Craig, J. V., Lancaster, G. A., Williamson, P. R., & Smyth, R. L. (2000). Temperature measured at the axilla compared with rectum in children and young people: Systematic review. *British Medical Journal*, *320*, 1174–1178.

Drake-Brockman, T. F., Hegarty, M., Chambers, N. A., & von Ungern-Sternberg, B. S. (2014). Monitoring temperature in children undergoing anaesthesia: A comparison of methods. *Anaesthesia and Intensive Care*, *42*, 315–320.

Evron, S., Weissman, A., Toivis, V., Shahaf, D. B., You, J., Sessler, D. I., et al. (2017). Evaluation of the Temple Touch Pro, a novel noninvasive core-temperature monitoring system. *Anesthesia and Analgesia*, *125*, 103–109.

Forrest, A. J., Juliano, M. L., Conley, S. P., Cronyn, P. D., McGlynn, A., & Auten, J. D. (2017). Temporal artery and axillary thermometry comparison with rectal thermometry in children presenting to the ED. *American Journal of Emergency Medicine*, *35*, 1855–1858.

Geijer, H., Udumyan, R., Lohse, G., & Nilsagard, Y. (2016). Temperature measurements with a temporal scanner: Systematic review and meta-analysis. *British Medical Journal Open*, *6*, e009509.

Greenes, D. S., & Fleisher, G. R. (2001). Accuracy of a noninvasive temporal artery thermometer for use in infants. *Archives of Pediatrics and Adolescent Medicine*, *155*, 376–381.

Greenes, D. S., & Fleisher, G. R. (2004). When body temperature changes, does rectal temperature lag? *Journal of Pediatrics*, *144*, 824–826.

Gunawan, M., Soetjiningsih, I., & Kardana, M. (2010). Comparison of the accuracy of body temperature measurements with temporal artery thermometer and axillary mercury thermometer in term newborns. *Paediatrica Indonesiana*, *50*, 67–72.

Hamilton, P. A., Marcos, L. S., & Secic, M. (2013). Performance of infrared ear and forehead thermometers: A comparative study in 205 febrile and afebrile children. *Journal of Clinical Nursing*, *22*, 2509–2518.

Hanneman, S. K. (2008). Design, analysis and interpretation of method-comparison studies. *AACN Advances in Critical Care*, *19*, 223–234.

Hebbar, K., Fortenberry, J. D., Rogers, K., Merritt, R., & Easley, K. (2005). Comparison of temporal artery thermometer to standard temperature measurements in pediatric intensive care unit patients. *Pediatric Critical Care Medicine*, *6*, 557–561.

Hoffman, R. J., Etwaru, K., Dreisinger, N., Khokhar, A., & Husk, G. (2013). Comparison of temporal artery thermometry and rectal thermometry in febrile pediatric emergency department patients. *Pediatric Emergency Care*, *29*, 301–304.

Holzhauer, J. K., Reith, V., Sawin, K. J., & Yen, K. (2009). Evaluation of temporal artery thermometry in children 3–36 months old. *Journal for Specialists in Pediatric Nursing*, *14*, 239–244.

Hooper, V. D., & Andrews, J. O. (2006). Accuracy of noninvasive core temperature measurement in acutely ill adults: The state of the science. *Biological Research for Nursing*, *8*, 24–34.

Hurwitz, B., Brown, J., & Altmiller, J. (2015). Improving pediatric temperature measurement in the ED. *American Journal of Nursing*, *115*, 48–55.

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

Kimberger, O., Cohen, D., Illievich, U., & Lenhardt, R. (2007). Temporal artery versus bladder thermometry during perioperative and intensive care unit monitoring. *Anesthesia and Analgesia*, *105*, 1042–1047.

Kottner, J., Audigé, L., Brorson, S., Donner, A., Gajewski, B. J., Hróbjartsson, A., et al. (2011). Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *Journal of Clinical Epidemiology*, *64*, 96–106.

Lawson, L., Bridges, E. J., Ballou, I., Eraker, R., Greco, S., Shively, J., et al. (2007). Accuracy and precision of noninvasive temperature measurement in adult intensive care patients. *American Journal of Critical Care*, *16*, 485–496.

Lee, G., Flannery-Bergey, D., Randall-Rollins, K., Curry, D., Rowe, S., & Teague, M. (2011). Accuracy of temporal artery thermometry in neonatal intensive care infants. *Advances in Neonatal Care*, *11*, 62–70.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gotzche, P. C., Ioannidis, J. P. A., et al. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: Explanation and elaboration. *Journal of Clinical Epidemiology*, *62*, e1–e34.

Mangat, J., Standley, T., Prevost, A., Vasconcelos, J., & White, P. (2010). A comparison of technologies used for estimation of body temperature. *Physiological Measurement*, *31*, 1105–1118.

Marable, K., Shaffer, L. E., Dizon, V., & Opalek, J. M. (2009). Temporal artery scanning falls short as a secondary, noninvasive thermometry method for trauma patients. *Journal of Trauma Nursing*, *16*, 41–47.

Mogensen, C. B., Wittenhoff, L., Fruerhoj, L., & Hansen, S. (2018). Forehead or ear temperature measurement cannot replace rectal measurements, except for screening purposes. *BMC Pediatrics*, *18*, 15.

Moore, A. H., Carrigan, J. D., Solomon, D. M., & Tart, R. C. (2015). Temporal artery thermometry to detect pediatric fever. *Clinical Nursing Research*, *24*, 556–563.

Nimah, M. M., Bshesh, K., Callahan, J. D., & Jacobs, B. R. (2006). Infrared tympanic thermometry in comparison with other temperature measurement techniques in febrile children. *Pediatric Critical Care Medicine*, *7*, 48–55.

Niven, D. J., Gaudet, J. E., Laupland, K. B., Mrklas, K. J., Roberts, D. J., & Stelfox, H. T. (2015). Accuracy of peripheral thermometers for estimating temperature. A systematic review and meta-analysis. *Annals of Internal Medicine*, *163*, 768–777.

Odinaka, K. K., Edelu, B. O., Nwolisa, C. E., Amamilo, I. B., & Okolo, S. N. (2014). Temporal artery thermometry in children younger than 5 years: A comparison with rectal thermometry. *Pediatric Emergency Care*, *30*, 867–870.

O'Grady, N. P., Barie, P. S., Bartlett, J. G., Bleck, T., Carroll, K., Kalil, A. C., et al. (2008). Guidelines for evaluation of new fever in critically ill adult patients. 2008 update from the American College of Critical Care Medicine and the Infectious Diseases Society of America. *Critical Care Medicine*, *36*, 1330–1349.

Opersteny, E., Anderson, H., Bates, J., Davenport, K., Husby, J., Myking, K., et al. (2017). Precision, sensitivity and patient preference of non-invasive thermometers in a pediatric surgical acute care setting. *Journal of Pediatric Nursing*, *35*, 36–41.

Paes, B. F., Vermeulen, K., Brohet, R. M., van der Ploeg, T., & de Winter, J. P. (2010). Accuracy of tympanic and infrared skin thermometers in children. *Archives of Disease in Childhood*, *95*, 974–978.

Penning, C., van der Linden, J. H., Tibboel, D., & Evenhuis, H. M. (2011). Is the temporal artery thermometer a reliable instrument for detecting fever in children? *Journal of Clinical Nursing*, *20*, 1632–1639.

Reitsma, J. B., Glas, A. S., Rutjes, A. W., Scholten, R. J., Bossuyt, P. M., & Zwinderman, A. H. (2005). Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology*, *58*, 982–990.

Reynolds, M., Bonham, L., Gueck, M., Hammond, K., Lowery, J., Redel, S., et al. (2014). Are temporal artery temperatures accurate enough to replace rectal temperature measurement in pediatric ED patients? *Journal of Emergency Nursing*, *40*, 46–50.

Sahin, S. H., Duran, R., Sut, N., Colak, A., Acunas, B., & Aksu, B. (2012). Comparison of temporal artery, nasopharyngeal, and axillary temperature measurement during anesthesia in children. *Journal of Clinical Anesthesia*, *24*, 647–651.

Schuh, S., Komar, L., Stephens, D., Chu, L., Read, S., & Allen, U. (2004). Comparison of the temporal artery and rectal thermometry in children in the emergency department. *Pediatric Emergency Care*, *20*, 736–741.

Selent, M. U., Molinari, N. M., Baxter, A., Nguyen, A. V., Siegelson, H., Brown, C. M., et al. (2013). Mass screening for fever in children: a comparison of 3 infrared thermal detection systems. *Pediatric Emergency Care*, *29*, 305–313.

Siberry, G. K., Diener-West, M., Schappell, E., & Karron, R. A. (2002). Comparison of temple temperatures with rectal temperatures in children under two years of age. *Clinical Pediatrics*, *41*, 405–414.

Suleman, M. I., Doufas, A. G., Akça, O., Ducharme, M., & Sessler, D. I. (2002). Insufficiency in a new temporal-artery thermometer for adult and pediatric patients. *Anesthesia and Analgesia*, *95*, 67–71.

Teran, C. G., Torrez-Llanos, J., Teran-Miranda, T. E., Balderrama, C., Shah, N. S., & Villarroel, P. (2012). Clinical accuracy of a non-contact infrared skin thermometer in paediatric practice. *Child: Care, Health and Development*, *38*, 471–476.

Tipton, E., & Shuster, J. (2017). A framework for the meta-analysis of Bland-Altman studies based on a limits of agreement approach. *Statistical Medicine*, *36*, 3621–3635.

Titus, M. O., Hulsey, T., Heckman, J., & Losek, J. D. (2009). Temporal artery thermometry utilization in pediatric emergency care. *Clinical Pediatrics*, *48*, 190–193.

Whiting, P. F., Rutjes, A. W., Westwood, M. E., Mallett, S., Deeks, J. J., Reitsma, J. B., et al. (2011). QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine*, *155*, 529–536.

Williamson, P. R., Lancaster, G. A., Craig, J. V., & Smyth, R. L. (2002). Meta-analysis of method comparison studies. *Statistics in Medicine*, *21*, 2013–2025.

Zhen, C., Xia, Z., Long, L., & Pu, Y. (2014). Accuracy of infrared ear thermometry in children: A meta-analysis and systematic review. *Clinical Pediatrics*, *53*, 1158–1165.