Received: 14 August 2020 | Revised: 30 December 2020 | Accepted: 31 January 2021

DOI: 10.1002/nop2.831

RESEARCH ARTICLE

NursingOpen Open Access WILEY

Threshold for defining fever varies with age, especially in children: A multi-site diagnostic accuracy study

Rajesh Kasbekar[1] | Aftab Naz[2] | Lorenzo Marcos[3] | Yingjie Liu[4] | Kristine Hendrickson[1] | James C. Gorsich[1] | Matt Baun[1]

[1]Kaz USA, Inc., a Helen of Troy Company, Marlborough, MA, USA

[2]Madera Family Medical Group, Madera, CA, USA

[3]Hospital del Niño Jesús, Tucumán, Argentina

[4]Merck Sharp & Dohme Corp, North Wales, PA, USA

Correspondence
Rajesh Kasbekar, Kaz USA, Inc., a Helen of Troy Company, Marlborough, MA, USA.
Email: RKasbekar@helenoftroy.com

Funding information
The study was funded by Kaz USA, Inc., a Helen of Troy company. No part of the study in terms of data collection or data analysis was influenced by Kaz USA, Inc., a Helen of Troy company as related to the integrity of the study.

Abstract

Aim: The American Academy of Pediatrics and the European Centre for Pediatric and Adolescent Medicine guideline define fever as a temperature >38.0°C for all ages and recommend use of rectal thermometers in children <3 years. Based on new literature, this definition of fever using a fixed threshold of 38.0°C needs to be re-examined.

Design: A multi-site diagnostic accuracy study was conducted to compare an "age-based" threshold model with a "fixed" threshold over 38.0°C on a total of 894 patients of which 373 were ill.

Methods: The "age-based" and "fixed" threshold fever determinations were then compared to a clinical categorization ("well" or "ill") conducted by a clinician through a comprehensive examination.

Results: The sensitivity and accuracy for the age-based thresholds were found to be superior to the fixed thresholds in all ages and current ear thermometers were found equivalent to rectal thermometers in infants <6 months.

KEYWORDS

adult, age-based, body temperature, children, fever, infant, newborn, paediatrics, parents, sensitivity, specificity, thermometer

1 | INTRODUCTION

The threshold between normal body temperature and fever has not been clear. The American Academy of Pediatrics (AAP) as well as the European Centre for Pediatric and Adolescent Medicine (ECPA) guideline defines fever as a temperature >100.4°*F* (38.0°C) [Wyckoff, 2009; Niehues, 2013] to be the current practice, regardless of age. They also recommend use of rectal thermometers in children <3 years. Several currently commercialized thermometers use such fixed-threshold-based fever alarms in their thermometers to inform the user of body temperature that is not normal.

Nonetheless, more recently several publications and organizations have challenged this fixed threshold between normal temperature and fever regardless of age, specifically within the first 3 years after birth. Several studies trying to correlate body temperature to illness and infection have concluded that the upper threshold of normal body temperature varies with age [Herzog et al., 2011]. The National Institute for Health and Care Excellence [NICE, 2019] has suggested that children between 0–3 months with a temperature

Equator Research Reporting Checklist: STARD.

of 38°C or higher are in a high-risk group for serious illness, while children aged 3–6 months with a temperature of 39°C or higher are in an intermediate-risk group for serious illness.

The rectal site has historically been the favourable measurement site for children, as it closely reflects core body temperature measurement compared to other sites and is less impacted by environmental temperature. However, it can be uncomfortable for children and carries a risk of contamination if not disinfected properly between uses [Batra et al., 2012].

Our objectives for this study were to test the hypothesis proposed by Herzog et al, that age-based fever thresholds for measured temperature are better estimators of illness and infection. A second objective was to compare state-of-the-art ear thermometers with rectal thermometers that are currently recommended for newborns and children under the age of six.

2 | BACKGROUND

Although the idea of age-stratified fever guidelines is not new to the medical community, an agreement on the specifics of these guidelines does not exist, even between medical professionals. Physicians continue to practice based on personal preferences of treatment and fever management. This lack of agreement among the medical community is cause for concern and these misconceptions, regarding fever and treatment, cause unnecessary aggressive and inappropriate management of feverish children. [Demir et al., 2012; Bettinelli et al., 2013]. Recently, novel diagnostic tools and new risk stratification tools such as the Step-by-Step approach and the Pediatric Emergency Care Applied Research Network prediction rule have been proposed to guide the management of febrile young infants in the emergency department [Palladino et al., 2019; Pade et al., 2019].

In addition, determination and management of fever needs to start at home, with informed parents. Parents need to be equipped with the knowledge (from the medical professional) when to care for their children at home and when further treatment is required. Historically, parents have a poor understanding of fever and little or no information about its beneficial role in diseases. [El-Radhi et al., 2012; Al-Eissa et al., 2000, Pusic et al., 2007, Schmitt et al., 1980]. There is a perceived need to improve recognition, assessment and management of fever with regards to underlying illness in children, furthering parental understanding of fever and fever management in relation to the age of the child [Demir et al., 2012; El-Radhi et al., 2012, Black et al., 2016].

To add to the confusion, there is a difference in temperature reading depending on the physiological site that is used to take the thermometer reading. A difference of 0.4°C exists between a rectal and an oral reading, while a difference of 0.7°C exists between a rectal and an axillary reading. The Canadian Pediatric Society has stated that a normal temperature range for a rectal reading is between 36.6°C and 38°C (97.9–100.4°F), while a normal temperature range for an oral reading is between 35.5°C and 37.5°C (95.9–99.5°F) [Canadian Pediatric Society, 2015].

Evidence is becoming more readily available regarding the validity and correlation between fever determination and age stratification [Herzog et al., 2011]. While several studies [Chiu et al., 1997; Pantell et al., 2004; Kuppermann et al., 1998; Hausfater et al., 2008; Laupland et al., 2008; Sund Lavender et al., 2002] have examined this correlation in specific age groups such as infants, toddlers, adults or the elderly, we have not seen any study to date that has compared the upper limits of normal temperature in healthy and febrile individuals of varying age groups using commercially available thermometers and comparing them with actual clinical diagnosis of "well" or "ill" performed by a healthcare professional using clinical impressions based on a comprehensive patient examination.

Several studies have addressed the accuracy of the ear thermometer compared to rectal with the goal of creating other reliable options for temperature measurement. One such study [Mogensen et al., 2018] examined over 900 patients, ages 0–18, using temporal, tympanic and rectal thermometers. The study found that the ear thermometer reached a sensitivity of over 90% in the 0.5–5-year age group with the area under the Receiver Operating Characteristic Curve (ROC) of 0.972 and a 95% confidence interval between 0.963–0.981. Another study [Nimah et al., 2006] concluded that ear thermometer measurements more accurately reflected core temperatures than any other measurement site during febrile and non-febrile periods in children and that ear measurements are a reproducible and relatively non-invasive substitute for bladder or rectal measurements in febrile children. A systematic review and network meta-analysis [Pecoraro et al., 2020] of over 12,000 patients in 46 studies showed that tympanic infrared thermometer measurements were not statistically different from zero when compared to rectal measurements. Pecoraro et al. also concluded that using a fixed fever cut-off temperature of 38°C (100.4°F), tympanic infrared thermometer measurements had a high sensitivity but a poor specificity using rectal measurements as the reference.

This study was designed, therefore, to address the key question of whether an age-based model of defining fever as suggested by Herzog et al [Herzog et al., 2011] has an improved clinical correlation with an actual clinical diagnosis of illness or infection when compared to a fixed fever threshold model. We also address a second important question of whether a well-designed infrared ear thermometer can be substituted for rectal thermometry, as is the current recommendation, in newborns as well as children under the age of 6 months.

Our critical assumption remains that fever should be considered as only one marker of potentially treatable illness and the overall impression of the patient's state determined by medical history and physical examination should be used to diagnose and manage illness using a risk stratification strategy [Avner et al.,2009; NICE, 2019; Palladino et al., 2019].

3 | DESIGN

In this prospective, two-arm clinical study, the determination of fever (using an age-based fever determination model programmed in a "state of the art" commercialized ear thermometer) was the primary endpoint in the 1st arm and was compared to the historical control of a "gold standard diagnosis for illness or infection" performed by a

healthcare professional (HCP) based on a clinical impression formed from a comprehensive patient examination. This diagnosis is considered as the absolute truth since it is based on a comprehensive examination of various diagnostic parameters and laboratory tests and the ideal reference for comparison to temperature measurement [Barbi et al., 2017]. The comparison was performed in five separate stratified age groups from birth to >65 years of age using a statistically valid sample size in each stratum. Another professional grade "ear or rectal" thermometer with fixed fever thresholds was used in the 2nd arm as the active control or the REFERENCE device. Sample size was calculated based on assumptions of type 1, type 2 error and prevalence of fever. The primary objective was to determine whether the sensitivity and specificity values estimated with this age-based fever determination model could be predicted to be above an "a priori acceptance criteria of 80 percent" with 95% confidence when compared to the diagnosis performed by a trained clinician, based on the clinical impression for the general population.

A secondary goal of the study was to compare a well-designed "ear" thermometer with age-based fever thresholds to a professional grade "rectal" thermometer with fixed fever thresholds for estimation of an "ill status" in infants <6 months. Therefore, the sensitivity, specificity and accuracy of the "test" ear thermometer obtained from the readings taken above were compared to the sensitivity, specificity as well as accuracy of a professional grade "rectal" thermometer with a fixed fever threshold for children <6 months (active control) when compared to the clinical diagnosis from the HCP (gold standard). Bias and Limits of Agreement between the test and REFERENCE device were also compared.

4 | METHOD[1]

4.1 | Sample size and subject selection

A sample size of over 600 subjects with at least 120 subjects in each age group strata had a power of well over 90% ($\beta = 0.1$) for multiplicity testing of the overall sensitivity and specificity simultaneously, assuming an alpha-level of 5% ($\alpha = 0.05$) and where the febrile prevalence is conjectured to be at least 30% [Pepe, 2003]. Our sample size requirement of over 850 was used to allow for attrition and for age strata as well as for formal statistical power calculations.

The study was divided into five age groups (Table 1). Participants were accepted into the study until an age category was filled with at least 120 patients in that age group and at least 30% and up to a maximum of 60% were febrile, having a temperature measurement of 37.5°C (99.5°F) or above, taken with a professional grade ear thermometer.

4.2 | Exclusion criteria

Subjects were excluded if they presented with acute life-threatening infections, anatomical abnormalities that would affect temperature or arterial blood pressure, pregnancy or lactation, ear diseases, abnormal skin conditions or scar tissue at the measurement site, hypothermia, those who were currently being medicated with antibiotics, analgesics or antipyretics administered <4 hr before enrolment or that were bedridden. Subjects were enrolled without regard to gender or ethnic background, or presence of other disabilities beyond those conditions described in the exclusion criteria above.

4.3 | Site selection

Subjects were recruited from six clinical sites across the United States and two international sites, one each in China and Argentina between December of 2014 and October of 2016. The clinical sites in the United States were the Jackson Clinic in Jackson Tennessee, Buena Salud Pediatrics in San Jose California, Yuma Regional Outpatient Pediatrics Clinic in Yuma Arizona, Middletown Family Care Associates in Middleton Delaware, Comprehensive Clinical Research in Berlin New Jersey and Madera Family Medical Group in Madera California.

TABLE 1 Age groups with corresponding sample size and ill/well subject population

IRT testing hypothesis group	Age subsets	Sample size	Ill	Well
0–3 months	Neonate birth to 1 month of age	79	23	56
	Infant >1–3 months of age	90	37	53
Over 3–36 months	Infant >3 months–2 years of age	176	84	92
	Child >2–3 years of age	79	55	24
Over 36 months to ≤12 years of age	Child >3–6 years of age	64	37	27
	Child >6–12 years of age	79	42	37
Over 12 years and adults ≤65 years of age	Adolescent >12–21 years of age	89	41	48
	Adult >21–65 years of age	96	29	67
Over 65 years	>65–85 years of age	81	17	64
	>85 years	61	8	53
All combined (Male)	0–90 years	894 (427 male)	373	521

The two international sites were No. 251 Hospital in Beijing in China and Hospital del Niño Jesús in Tucumán in Argentina.

4.4 | Ethics

Approval for the six sites in the United States was obtained by a central Institutional Review Board (IRB), the Alpha IRB out of California. Approvals for the site at No. 251 Hospital in Beijing (China) and at the Hospital del Niño Jesús in Tucumán (Argentina) were granted by the Human Subjects Protection Review Boards of the individual hospital prior to subject enrolment. At each site, the recruiting nurse ascertained from the adults or from parents or guardians whether the children were interested in participating. For every site, inclusion in the study required written, signed informed consent for the adults from the parent or guardian and an assent form or verbal assent for children 7–17 years old.

4.5 | Thermometers

The thermometer investigated in this study was the ThermoScan® Braun IRT6520 infrared ear thermometer designated as IRT (Helen of Troy, Inc) which measures the temperature in the ear canal but provides an oral equivalent reading. The Age Precision™ feature, present in the IRT, has age-based cut-off points for fever (Table 2). This feature gives an indicator light warning of no light (white) for normal or sub-normal temperature, yellow for elevated and red for high temperature depending on the temperature reading and the age-based fever cut-off points. All yellow and red backlights were considered as a febrile output.

The SureTemp® Plus electronic contact thermometer (Welch Allyn), designated as SURETEMP was used in monitor mode and was designated as the professional grade rectal REFERENCE thermometer for children under 6 months. The SURETEMP in monitor mode was chosen because taking a 3-min reading with SURETEMP in monitor mode ensures that any bias that might be introduced due to an algorithm will not influence the reading, which is a direct reflection of measurement by a calibrated thermistor. The SURETEMP has been used as a reference standard in other thermometer comparisons [Giuliano et al., 2000; Mangat et al., 2010].

The professional version of the Braun Infrared ear thermometer, the ThermoScan® PRO 4000 ear thermometer, designated as PRO, was used as the REFERENCE thermometer for subjects ≥6 months. The PRO ear thermometer has been widely accepted as the standard of care in infrared ear thermometry [Hamilton et al., 2013]. The PRO and the SURETEMP are collectively referred to as the "REFERENCE."

Prior to the clinical study, all thermometers came calibrated by their respective manufacturers in their individual laboratory settings to assure reliability and laboratory accuracy. Prior to subject enrolment each day, ear thermometers were verified in their respective calibration devices to assure that the thermometers were

TABLE 2 Age stratification limits

Site/Display	Age range	Sub-normal body temperature (White or no backlight)	Range of normal body temperature (White or no backlight)	Fever (Yellow)	High fever (Red)
Site measured: Ear Display: Oral equivalent reading	≤3 months	<96.4°F or <35.8°C	≥96.4°F or ≥35.8°C ≤99.4°F or ≤37.4°C	N/A	>99.4°F or >37.4°C
	>3 months ≤36 months	<95.7°F or <35.4°C	≥95.7°F or ≥35.4°C ≤99.6°F or ≤37.6°C	>99.6°F or >37.6°C ≤101.3°F or ≤38.5°C	>101.3°F or >38.5°C
	>36 months	<95.7°F or < 35.4°C	≥95.7°F or ≥35.4°C ≤99.9°F or ≤ 37.7°C	>99.9°F or >37.7°C ≤103.0°F or ≤39.4°C	>103.0°F or >39.4°C

Note: The IRT (test article) is an ear thermometer that measures in the ear but outputs an oral equivalent reading. Hence oral age stratification limits were used by the Age Precision™ feature in the IRT to estimate fever and are displayed above.

functioning properly and measuring within specifications. The professional grade rectal thermometers (monitor mode) were also checked for proper calibration prior to study initiation.

For children younger than 6 months of age, the fixed threshold of febrile for a rectal reading for this study was a temperature ≥38.0°C (100.4°F) [Wyckoff, 2009; Niehues, 2013; European Committee for Standardization, 2003] measured by the professional grade rectal thermometer in monitor mode for at least 3 min. For children older than 6 months of age, the fixed threshold of febrile for this study was >38.0°C [Wyckoff, 2009; European Committee for Standardization, 2003] measured by the professional grade ear thermometer. Prior to taking measurements, the area for taking measurements was prepared in accordance with the manufacturer's instructions.

All five temperature measurements with each test thermometer and one temperature measurement with the reference thermometer were obtained in the presence of a physician observer by a HCP, as recommended by the European Committee for Standardization [European Committee for Standardization, 2003]. The HCP was well trained on all equipment and had years of experience taking thousands of temperature measurements. All ear measurements (unless contraindicated) were measured in the left ear. There was at least a 1-min wait time between temperature measurements.

All equipment was thoroughly cleaned according to the manufacturers' instructions at the end of each workday. Appropriate manufacturer probe covers were used and were changed for each measurement for each subject. SURETEMP rectal probes with probe covers were used for all rectal measurements. Probe covers manufactured by Braun were used on both the IRT and PRO ear thermometers.

4.6 | Randomization

Each site was supplied with three thermometers of each type. The selection of the IRT and the REFERENCE thermometers as well as the sequence of temperature measurements were randomized using a pre-determined randomization chart consisting of a random allocation sequence generated by an independent biostatistician. The categorization of the subjects as ill or well was done prior to the use of the devices. The site investigator at each site enrolled the subjects, and the site coordinator for that site assigned the device to each subject. The HCP taking the measurements reviewed the randomization assignment just before the measurement started, and there were no exclusions after the randomization. All measurements were taken in accordance with the training in a similar manner.

4.7 | Patient examination

Determination of the objective "well" or "ill" status was based on clinical impression, diagnosis and decision to use interventions and order tests. The "moderately or mildly ill" category included patients with mild forms of illnesses; categories of illness are defined (Table 3) over the entire population as well as across all age groups. The analysis of "moderately ill" and "ill" were both treated as an indication of an "ill" status. The use of clinical examination and prudent utilization of laboratory tests were considered as the cornerstone of safe management of febrile children and the ideal reference for comparison with temperature measurement [Barbi et al., 2017; Nijman et al., 2013].

4.8 | Statistical methods

The primary endpoint of the study was to compare the sensitivity and specificity of the fever determination by the Age Precision™ feature in the IRT (Test article) ear thermometer with the fever determination of "ill" and "well" patients, as determined by a clinical impression on the overall population comprising patients ranging in age from neonates to >85 years as well as by stratified age categories (Table 1). Using point estimates and its 95% confidence interval, an acceptance-criteria of 80% for both sensitivity and specificity [Van den Bruel, 2011] in the overall population was used for the lower limit of the 95% confidence interval to show the effectiveness of this method. Van deb Bruel et al. found the sensitivity and specificity of the most accurate "predictive" laboratory test when comparing illness and infection to be 75.1% and 76.1% respectively. Using a fudge factor, we used an acceptance-criteria of 80% as a conservative estimate for the lower limit of the 95% confidence interval for sensitivity and specificity.

In addition, a post hoc analysis was conducted to compare the sensitivity, specificity and accuracy values of the IRT with the sensitivity, specificity and accuracy values of the REFERENCE in comparison to the clinical impression of "ill" and "well" patients in the overall population as well as in each of the age-stratified categories.

The validity of this comparison assumes that the thermometers used in the study agree well with each other in terms of bias and standard deviation and have a high level of accuracy in terms of state of the art in thermometry. Therefore, bias and standard deviation were calculated and a Bland–Altman analysis conducted to compare the agreement between the devices in terms of the bias and standard deviation, as recommended by the International Standardization Organization (International Organization for Standardization, 2017) and American Society for Testing and Materials E1965 (ASTM International, 2016).

4.9 | Sensitivity, specificity, accuracy and receiver operating characteristic (ROC) analysis

Sensitivity and specificity were calculated for the overall population as well as for each of the stratified age groups. Ninety-five (95) % confidence intervals were then calculated assuming normal distribution of the sampling distribution. Similarly, accuracy was calculated for the overall population as well as for each of the stratified age

TABLE 3 Definition of ill or well

Age group	Seriously ill (patient categorized as "seriously ill" if he/she had any of the following)	Moderately/Mildly ill (patient categorized as "moderately/mildly ill" if he/she had any of the following and did not have any symptom or intervention as those for "seriously ill")	Well (patient categorized as "well" if he/she did not have any symptom or intervention as those for "seriously ill" and "moderately ill")
Symptoms			
Birth to 1 month			
	Jaundice	Mild jaundice	No notable symptoms
	Vomiting or diarrhoea, bloody diarrhoea	Spitting up	
	Reduced movements		
	Reduced sucking		
	Excessive crying or high-pitched cry or grunt		
	Cyanosis		
	Dyspnoea, tachypnoea, apnoea or gasping		
	Chest retractions		
	Excessive irritability		
	Neck retraction		
	Blank look		
	Bulging fontanelle		
	Seizures		
	Abnormal heart rate		
	Swollen belly area		
Symptoms			
>1–3 months			
	Severe dehydration (e.g. dry diaper more than 12 hr, sunken soft spot on skull)	Moderate dehydration (e.g. dry diapers)	No notable symptoms
	Refusing milk	Decreased milk intake	
	Vomiting or diarrhoea repeatedly, bloody diarrhoea	Moderate vomiting or diarrhoea	
	Excessive irritability	Higher irritability than usual	
	Seizures	Reduced movements	
	Excessive cry or high-pitched cry or grunt	Excessive crying	
	Cyanosis	Sweating	
	Chest retractions		
	Neck retraction		
	Dyspnoea		
	Blank look		
	Bulging fontanelle		
	Abnormal heart rate		
	Swollen belly area		
Symptoms			
>3 months–2 years			
	Severe dehydration (e.g. dry diaper more than 12 hr, sunken soft spot on skull)	Moderate dehydration (e.g. dry diapers)	No notable symptoms
	Refusing to eat or drink	Decreased appetite	
	Severe shivering	Moderate shivering	

(Continues)

TABLE 3 (Continued)

	Vomiting or diarrhoea repeatedly, bloody diarrhoea	Moderate vomiting or diarrhoea	
	Excessive irritability	Higher irritability than usual	
	Seizures	Reduced movements	
	Excessive cry or high-pitched cry or grunt	Excessive crying	
	Cyanosis	Sweating	
	Chest retractions	Red rash	
	Neck retraction		
	Dyspnoea		
	Rash with purple (blood-coloured) spots or dots		
	Abnormal heart rate		
	Swollen belly area		
	Other evidence of severe infection (e.g. deep tissue injury)		
Symptoms			
>2–3 years			
	Severe dehydration (e.g. no urine in >12 hr, sunken eyes, tented skin, very dry mouth, etc.)	Moderate dehydration (e.g. dry mouth or mucous membranes, etc.)	No notable symptoms
	No appetite	Decreased appetite	
	Severe shivering	Moderate shivering	
	Vomiting or diarrhoea repeatedly, bloody diarrhoea	Moderate vomiting or diarrhoea	
	Very weak (e.g. inability to stand or perform normal activities)	Fatigue or decreased activity	
	Seizures	Reduced movements	
	Severe throat swelling	Sore throat	
	Cyanosis	Sweating	
	Chest retractions	Red rash	
	Neck retraction		
	Severe dyspnoea (e.g. struggling for each breath, unable to speak)		
	Rash with purple (blood-coloured) spots or dots		
	Abnormal heart rate		
	Unexplained irritability		
	Other evidence of severe infection (e.g. deep tissue injury)		
Symptoms			
>3–6 years			
	Severe dehydration (e.g. no urine in >12 hr, sunken eyes, tented skin, very dry mouth, etc.)	Moderate dehydration (e.g. dry mouth or mucous membranes, etc.)	No notable symptoms
	No appetite	Decreased appetite	
	Severe shivering	Moderate shivering	
	Vomiting or diarrhoea repeatedly, bloody diarrhoea	Moderate vomiting or diarrhoea	
	Very weak (e.g. inability to stand or perform normal activities)	Fatigue or decreased activity	
	Seizures	Reduced movements	
	Severe throat swelling	Sore throat	
	Cyanosis	Sweating	
	Chest retractions	Red rash	
	Neck retraction		

(Continues)

TABLE 3 (Continued)

	Severe dyspnoea (e.g. struggling for each breath, unable to speak)		
	Rash with purple (blood-coloured) spots or dots		
	Abnormal heart rate		
	Unexplained irritability		
	Other evidence of severe infection (e.g. deep tissue injury)		
Symptoms			
>6–12 years			
	Severe dehydration (e.g. no urine in >12 hr, sunken eyes, tented skin, very dry mouth, etc.)	Moderate dehydration (e.g. dry mouth or mucous membranes)	No notable symptoms
	No appetite	Decreased appetite	
	Severe shivering	Moderate shivering	
	Vomiting or diarrhoea repeatedly, bloody diarrhoea	Moderate vomiting or diarrhoea	
	Very weak (e.g. inability to stand or perform normal activities)	Fatigue or decreased activity	
	Severe headache	Moderate headache	
	Severe muscle pain	Moderate muscle pain	
	Seizures or convulsions	Reduced movements	
	Severe throat swelling	Sore throat	
	Cyanosis	Sweating	
	Pain when urinating	Increased frequency/urgency to urinate	
	Chest retractions	Red rash	
	Neck retraction		
	Severe dyspnoea (e.g. struggling for each breath, unable to speak)		
	Rash with purple (blood-coloured) spots or dots		
	Stiff neck and pain when bending head forward		
	Unusual sensitivity to bright light		
	Abnormal heart rate		
	Mental confusion or difficult to awaken		
	Hallucinations		
	Unexplained irritability		
	Other evidence of severe infection (e.g. deep tissue injury)		
Symptoms			
>12–21 years			
	Severe dehydration (e.g. sunken eyes, tented skin and very dry mouth)	Moderate dehydration (e.g. dry mouth or mucous membranes and lightheaded)	No notable symptoms
	No appetite	Decreased appetite	
	Severe shivering	Moderate shivering	
	Vomiting or diarrhoea repeatedly, bloody diarrhoea	Moderate vomiting or diarrhoea	
	Very weak (e.g. inability to stand or perform normal activities)	Fatigue or decreased activity	
	Severe headache	Moderate headache	
	Severe muscle pain	Moderate muscle pain	

(Continues)

TABLE 3 (Continued)

	Seizures or convulsions	Reduced movements	
	Severe throat swelling	Sore throat	
	Cyanosis	Sweating	
	Pain when urinating	Increased frequency/urgency to urinate	
	Chest retractions	Red rash	
	Neck retraction		
	Severe dyspnoea (e.g. struggling for each breath, unable to speak)		
	Rash with purple (blood-coloured) spots or dots		
	Stiff neck and pain when bending head forward		
	Unusual sensitivity to bright light		
	Abnormal heart rate		
	Mental confusion or difficult to awaken		
	Hallucinations		
	Unexplained irritability		
	Other evidence of severe infection (e.g. deep tissue injury)		
Symptoms			
>21 years			
	Severe dehydration (e.g. sunken eyes, tented skin and very dry mouth)	Moderate dehydration (e.g. dry mouth or mucous membranes and lightheaded)	No notable symptoms
	No appetite	Decreased appetite	
	Severe shivering	Moderate shivering	
	Vomiting or diarrhoea repeatedly, bloody diarrhoea	Moderate vomiting or diarrhoea	
	Very weak (e.g. inability to stand or perform normal activities)	Fatigue or decreased activity	
	Severe headache	Moderate headache	
	Severe muscle pain	Moderate muscle pain	
	Seizures or convulsions	Reduced movements	
	Severe throat swelling	Sore throat	
	Cyanosis	Sweating	
	Pain when urinating	Increased frequency/urgency to urinate	
	Chest retractions	Red rash	
	Neck retraction		
	Severe dyspnoea (e.g. struggling for each breath, unable to speak)		
	Rash with purple (blood-coloured) spots or dots		
	Stiff neck and pain when bending head forward		
	Unusual sensitivity to bright light		
	Abnormal heart rate		
	Mental confusion or difficult to awaken		
	Hallucinations		
	Unexplained irritability		
	Other evidence of severe infection (e.g. deep tissue injury)		
Intervention			
All ages			
	Hospitalization	Chest X-ray	No intervention

groups. A receiver operating characteristic (ROC) was then used to compare the sensitivities and specificities of each strata.

A similar analysis was performed post hoc on the REFERENCE thermometers with the aim of comparing their relative performance using the fixed fever thresholds as recommended by the AAP. This analysis was performed on the overall population as well as on the stratified age groups.

5 | RESULTS

5.1 | Population summary

Eight hundred and ninety-five subjects 0–90 years were enrolled and completed the study. They were divided into five age groups with more than 120 in each age group. One participant was excluded from the final analysis due to missing data. There were no indeterminate measurements in the study. Of the 894 remaining participants, 567 (63%) were children under the age of 12, 373 (42%) were ill based on the patient impression of "ill" and 427 (48%) were male (Table 1). There were no adverse events identified as part of this study.

5.2 | Overall population

The accuracy, sensitivity and specificity values as well as the 95% confidence intervals for the overall population are shown in Table 4. The IRT thermometer predictions for illness were thus found to successfully meet the acceptance criteria for both the sensitivity and specificity in the overall population.

5.3 | Neonates and other age categories

The accuracy, sensitivity and specificity values for each age category are shown in Table 4 as well as in Figure 1 showing accuracy and Figure 2 showing sensitivity and specificity. The 95% confidence intervals for sensitivity and specificity met the stipulated 80% acceptance criteria in the 0–3 month, 3–36 month and other age categories. A receiver operating characteristic curve (ROC) showing the sensitivity and specificity trade-off for the TEST and REFERENCE devices was computed and is shown in Figure 3.

5.4 | Bland–Altman analysis between test and reference thermometers

In order to verify the key assumptions of the sensitivity and specificity analysis, Bland–Altman analysis (Bland & Altman 1986) was used to compare the IRT temperature readings to the SURETEMP rectal readings, using the first reading, for neonates 0–6 months as well as to the PRO ear readings for subjects >6 months (Figure 4). The average bias (difference) for overall comparisons was −0.02°C ($p < .001$) and the standard deviation of the bias was 0.30°C with the 95% confidence interval between −0.04°C and 0°C. The correlation coefficient for the IRT and the REFERENCE thermometers (SURETEMP + PRO) taken together was 0.94 ($p < .00$). This indicated a good agreement between the IRT ear readings and the SURETEMP rectal readings as well as between the IRT and PRO ear readings.

5.5 | "Ear thermometer readings and illness" versus "rectal thermometer readings and illness"

Rectal temperature measurement has been recommended by the AAP and the ECPA for temperature measurement in children under the age of 3 years. [Wyckoff, 2009; Niehues, 2013].

In the neonate age category of 0–3 months, the accuracy using the age-based fever threshold in the IRT ear thermometer was 93.5% compared to an accuracy of 72.8% using the fixed fever thresholds in the SURETEMP rectal thermometer. The sensitivity of the IRT was also substantially higher than the SURETEMP, while the specificity was similar in this age group. This clearly shows that the accuracy and sensitivity of the state-of-the-art "ear" thermometer (IRT) with age-based thresholds were substantially higher than the accuracy and sensitivity of the state-of-the-art "rectal" thermometer (SURETEMP) using fixed fever thresholds when predicting illness.

6 | DISCUSSION

Fever is an important prognosticator of illness and infections. It is a symptom of mostly viral infections but on occasion bacterial infections that can be treated with antibiotics. The likelihood of an infection increases generally with the magnitude of the fever, and therefore when defining a fever threshold, *the ideal fever threshold needs to be based on its correlation to actual illness or infection*. The optimal fever threshold selection should have sensitivity and specificity values that are as high as possible on the receiving operating characteristic (ROC) curve, so that the area under the curve is maximized. The fixed threshold of 100.4°*F* (38.0°C) provided as a guideline by the AAP and the ECPA, while reasonably safe, provides a poor prognostication to infections and illness as seen from the data in this study. Herzog et al [Herzog et al., 2011] took a different view and did an extensive review of the literature to study the different cut-off points and identify the lower limit of "fever" and "high fever" based on the patient's ages. Several other researchers [Chiu et al., 1997; Crain et al., 1982; Pantell et al., 2004] have concluded that the lower fever limit for children between birth and 3 months should be 100.4°*F* (38.0°C) when measured rectally. They also concluded that any elevated temperature above this threshold has high likelihood of infection and should be considered as high fever. Teach and Fleisher et al [Teach et al., 1997] and Kupperman et al [Kuppermann et al., 1998] have provided evidence and concluded that the lower limit for high fever

TABLE 4 Sensitivity, specificity and accuracy (overall cohort and by age stratum)

Age Strata	Parameter	IRT (Test article) Estimate *N*/m (%)	IRT (Test article) 95% Confidence interval	SURETEMP (reference article) Febrile >38.0°C Estimate *N*/m (%)	SURETEMP (reference article) Febrile >38.0°C 95% Confidence interval	PRO (reference article) Febrile >38.0°C Estimate *N*/m (%)	PRO (reference article) Febrile >38.0°C 95% Confidence interval	SURETEMP + PRO Combined Estimate *N*/m(%)	SURETEMP + PRO Combined 95% confidence
All	Sensitivity	317/373 (84.99%)	80.95%–88.46%	26/85 (30.59%)	21.05%–41.53%	124/288 (43.06%)	37.26%–48.99%	150/373 (40.21%)	35.20%–45.39%
	Specificity	513/521 (98.46%)	97.00%–99.33%	155/155 (100%)	97.65%–100.00%	360/366 (98.36%)	96.47%–99.40%	515/521 (98.85%)	97.51%–99.58%
	Accuracy	830/894 (92.84%)	90.95%–94.44%	181/240 (75.42%)	69.47%–80.73%	484/654 (74%)	70.46%–77.33%	665/894 (74.4%)	71.4%–77.2%
	Positive predictive value	317/325(97.54%)	95.21%–98.75%)	26/26 (100%)	100%–100%	124/130 (95.38%)	90.24%–97.88%	150/156 (96.15%)	91.79%–98.2%
	Negative predictive value	513/569(90.16%)	87.80%–92.10%	155/214 (72.43%)	69.52%–75.16%	360/524 (68.70%)	66.48%–70.84%	515/738 (69.8%)	68%–71.5%

Age strata	Parameter	Test article Age-based febrile determination Estimate *N*/m (%)	Test article Age-based febrile determination 95% Confidence interval	Reference article Febrile >38.0°C Estimate *N*/m (%)	Reference article Febrile >38.0°C 95% Confidence interval
0–3 months	Sensitivity	50/60(83.3%)	(71.5%, 91.7%)	14/60 (23.33%) (Rectal)	(13.4%, 36.0%)
	Specificity	108/109 (99.1%)	(95%, 99.98%)	109/109 (100%) (Rectal)	(96.7%, 100.0%)
	Accuracy	158/169 (93.49%)	88.65%, 96.71%)	123/169(72.8%)	(65.41%, 79.3%)
Over 3–36 months	Sensitivity	115/139 (82.7%)	(75.4%, 88.6%)	12/25 (48.0%) (Rectal) 47/114 (41.2%) (Oral) 59/139(42.45%) (Total)	(27.8%, 68.7%) (32.1%, 50.8%) (34.11%, 51.11%)
	Specificity	114/116 (98.28%)	(93.9%, 99.8%)	46/46 (100.0%) (Rectal) 68/70 (97.1%) (Oral) 114/116 (98.3%) (Total)	(92.3%, 100.0%) (90.06%, 99.65%) (93.91%, 99.79%)
	Accuracy	229/255 (89.8%)	(85.4%, 93.23%)	173/255 (67.8%) (Total)	(61.73%, 73.53%)

(Continues)

TABLE 4 (Continued)

Age strata	Parameter	Test article Age-based febrile determination		Reference article Febrile >38.0°C	
		Estimate N/m (%)	95% Confidence interval	Estimate N/m (%)	95% Confidence interval
Over 36 months to ≤12 years	Sensitivity	67/79 (84.81%)	(74.97%, 91.90%)	45/79 (56.9%) (Oral)	(45.3%, 68.1%)
	Specificity	61/64 (95.31%)	(86.91%, 99.02%)	61/64 (95.3%) (Oral)	(86.9%, 99.0%)
	Accuracy	128/143 (89.5%)	(83.29%, 94.01%)	106/143 (74.1%)	(66.14%, 81.08%)
Over 12 years and adults ≤65 years	Sensitivity	62/70 (88.57%)	(78.72%, 94.93%)	21/70 (30.0%) (Oral)	(19.6%, 42.1%)
	Specificity	114/115 (99.13%)	(95.25%, 99.98%)	115/115 (100.0%) (Oral)	(96.8%, 100.00%)
	Accuracy	176/185(95.14%)	(90.97%, 97.75%)	136/185 (73.5%)	(66.54%, 79.72%)
Over 65 years	Sensitivity	23/25 (92%)	(73.9%, 99.02%)	11/25 (44%) (Oral)	(24.4%, 65.1%)
	Specificity	116/117 (99.15%)	(91.37%, 99.98%)	116/117 (99.2%) (Oral)	(95.3%, 99.9%)
	Accuracy	139/142 (97.9%)	(94%, 99.57%)	127/142 (89.4%)	(83.18%, 93.97%)

Note: Sensitivity, specificity and accuracy are defined as follows:

Sensitivity = TP/(TP + FN).

Specificity = TN/(TN + FP).

Accuracy = (TP + TN)/(TP + TN+FP + FN).

Positive predictive value = TP/(TP + FP).

Negative predictive value = TN/(TN + FN).

in children between 3 months up to 36 months is 102.2°*F* (39.0°C) when measured rectally. Shaw et al [Shaw et al., 1998] have concluded that a threshold of 100.9°*F* (38.3°C) when measured rectally should be used as a lower fever limit in children between 3 months and 2 years. Hausfater et al [Hausfater et al., 2008] concluded, based on evidence generated using a tympanic thermometer, that the lower fever threshold limit for subjects between 6 years and 103 years should be 100.4°*F* (38.0°C), while Laupland et al [Laupland et al., 2008] also concluded that the lower fever threshold limit for "high fever" in subjects over 18 years should be 103.1°*F* (39.5°C). The normal body temperature in over 27 studies and thousands of adults between 1935 and 1999 was studied by Sund Lavender et al [Sund Lavender et al., 2002; Waalen et al., 2011] and they concluded that age might be a major factor in its connection to declining normal body temperature.

A synthesis of these and several other studies were used by Herzog et al [Herzog et al., 2011] to deduce an age-based model that can provide a better prognostication to illness in comparison to the fixed threshold model recommended by the AAP or the ECPA. The main objective of our study, therefore, was to validate this hypothesis proposed by Herzog et al [Herzog et al., 2011] using a diagnostic, well-powered clinical study. Our primary objective was to test this hypothesis and provide evidence that fever as defined by "age-based" thresholds and as calculated by the Age Precision™ feature of the IRT had a sensitivity and specificity more than or equal to 80% in the overall cohort. Post hoc analysis compared the accuracy, sensitivity and the specificity of this feature to the accuracy, sensitivity as well as the specificity of professional REFERENCE thermometers with fixed age thresholds as recommended by AAP and ECPA. One key assumption of this study was that the IRT had good agreement (adjusted for site-specific differences) with the REFERENCE thermometers.

As outlined in the results section above (Table 4), the values for sensitivity, specificity and accuracy meet the acceptance criteria. While the sensitivity values were found to be much lower using a

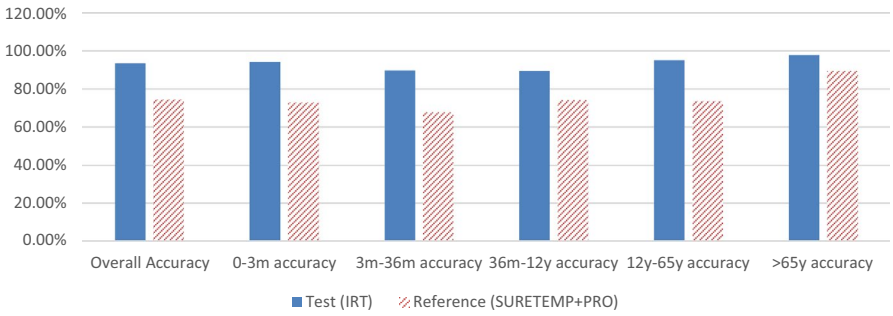

FIGURE 1 Comparison of accuracy for fever determination: Overall and Each Age Strata—Test versus Reference

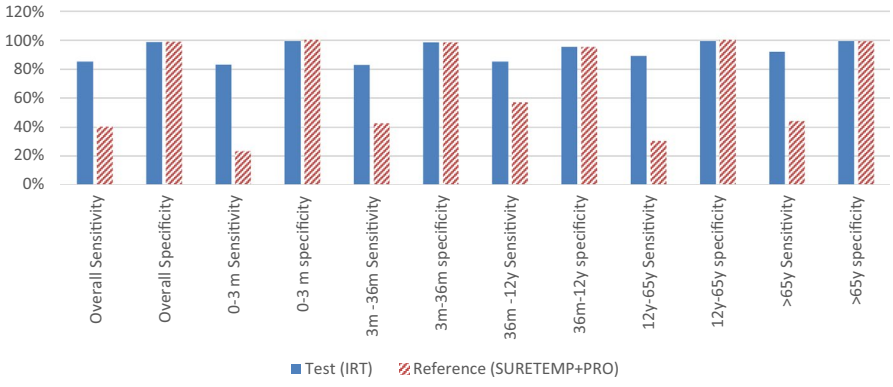

FIGURE 2 Comparison of sensitivity and specificity for fever determination: Overall and Each Age Strata—Test versus Reference

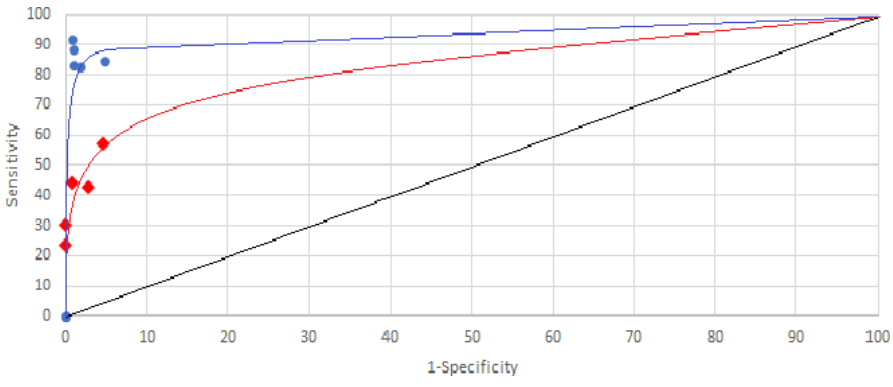

FIGURE 3 Receiver Operating Curve Test Device (in Blue) versus Reference Device (in Red)

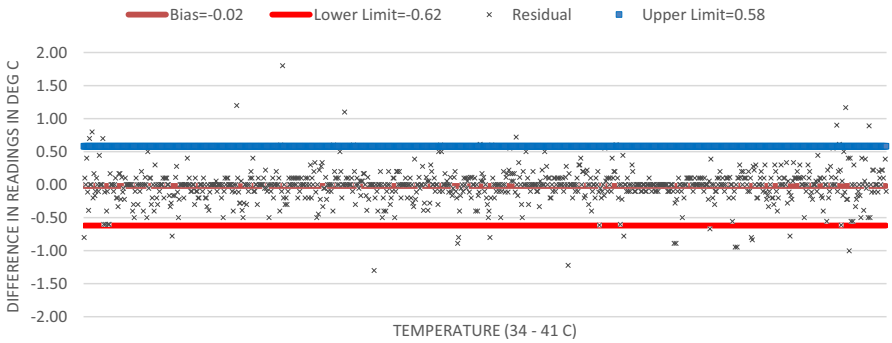

FIGURE 4 Bland–Altman Chart—Overall

fixed threshold of 38°C (100.4°F) as compared to the sensitivity values using the age-based fever thresholds in the IRT; the specificity values were comparable. This is consistent with the findings of Pecoraro et al [Pecoraro et al., 2020] who conducted a meta-analysis of 46 studies which compared infrared and digital thermometer measurements with rectal measurements taken by a "mercury in tube" or a "digital thermometer" as a reference. Using 38°C (100.4°F) as cut-off temperature, they found that tympanic infrared thermometry had a sensitivity of 0.77 (0.60, 0.88) and a specificity of 0.98 (0.95, 0.99). Figure 3 shows the Receiver Operating Characteristic curve (ROC) for the IRT and the REFERENCE thermometer. Clearly, the curve for the test thermometer shows a superior profile compared to the REFERENCE due to the improvement in sensitivity using the age-based model.

Bland–Altman analyses, as elaborated in the results section above, indicates that the difference in prediction for all instruments is in a narrow range. Both the bias and the standard deviations between the IRT ear thermometer and the rectal SURETEMP thermometer as well as the PRO ear thermometer is relatively small for this type of comparison. This is consistent with the observations of Nimah et al [Nimah et al., 2006].

6.1 | Limitations

The study was limited to the use of popular and widely used infrared ear thermometers, the IRT and the PRO that use a thermopile based infrared sensor and the popular thermistor based rectal thermometer, the SURETEMP, in lieu of invasive "true" core temperature measurements that are not practical.

We did not consider the effect of time of the day or sex of the subject on temperature variability, because we assumed that it would have a minimal effect on the fever thresholds. Temperatures tend to be lower in the morning compared to those in the evening and can also vary depending on whether the subject is male or female [Herzog et al., 2011]. Future work should include segmentation of the subject's sex and time of the day when the temperature was taken to see if it has any significant effect on the prognostication of illness.

To eliminate a clinical site bias, several clinical sites were used to conduct the study. While every effort was made to standardize the protocol (for estimation of the clinical impression for determining ill, moderately ill or well) and technique for temperature measurement, site and physician-based differences [Greene et al., 1981] could contribute to some variability in forming clinical impressions of illness and techniques used in taking a temperature reading.

6.2 | Relevance to clinical practice

Bacterial and viral infections cause illness and are cause for concern among physicians, especially in infants and toddlers. Fever status, as determined by a thermometer, is used not only as one parameter for forming a clinical impression of illness, but also for prescribing antibiotics and antipyretics.

An improvement in the determination of illness using a "fever thermometer" can not only improve timely medical intervention but also reduce unnecessary visits to a doctor's office. Several other publications [Pusnik et al., 2009; Baraff et al., 1993] provide further evidence that the management of young febrile children needs to be revisited and restructured to minimize the likelihood of unfavourable outcomes.

This new model will enable nurses to reduce false negatives when screening patients in the emergency room (using temperature measurements) or when prescribing antipyretics during phone consultations, for illness or infection. Healthcare professionals including nurses and physicians can program the site-specific "age-based thresholds" recommended by this model [Herzog et al., 2011] into their vital sign monitors that measure temperature. This will reduce the false negatives from alarms during monitoring and discharge of sick patients in hospitals and other healthcare institutions, based on a direct clinical correlation with illness or infection.

The framework provided by this study will also allow parents and caregivers to make a better choice in seeking medical intervention [Al-Eissa et al., 2000]. It will allow for initial management of fever at home, instead of unnecessary medical visits. In turn, this can help reduce medical costs for parents and reduce the waste of medical resources. As an example, other authors [Baker et al., 1993] have demonstrated a significant cost saving due to accurate prognosis of infants having illness or infection.

A thermometer, along with its associated features for defining fever, should yield low false negative and false positive readings

when compared against actual illness, especially in children under the age of five. There is continued scepticism on use of infrared ear thermometers in infants and toddlers as a substitute for thermistor based rectal thermometers. Based on the overwhelming amount of data comparing a state-of-the-art ear thermometer to a widely used commercial rectal thermometer with fixed fever thresholds as defined by the AAP, these data confirm that a well-designed infrared ear thermometer can also be recommended for newborns as well as children under the age of 6 months, in place of invasive rectal thermometry.

7 | CONCLUSION

In conclusion, our data clearly indicated that this empirical model for age-based fever thresholds suggested by Herzog et al [Herzog et al., 2011] showed a closer agreement (in terms of sensitivity and accuracy) between fever as a result of elevated temperatures and illness as identified by a clinical impression from a HCP. It clearly showed that temperature readings from a state-of-the-art "ear" thermometer with age-based model have an improved agreement with clinical impressions of illness as determined by an HCP, when compared to measurements taken by a professional "rectal" thermometer, especially in newborns and children under the age of 6 months.

ACKNOWLEDGEMENTS

We are grateful to the following people: Study Design: The study design was conducted by Rajesh Kasbekar and Yingjie Liu in collaboration with the US Food and Drug Administration, Centre for Devices and Radiological Health. Sample size calculations, power assumptions and statistical methodology were designed by Michelle Secic (M.S Biostatistics) from Secic Consulting. Data Collection: Data collection was done by the following healthcare professionals and their affiliations: Dr. Lorenzo Marcos: Hospital del Niño Jesús, Tucumán, Argentina. Dr. Yuzhi Zuo: No. 251 Hospital, Beijing, China. Dr. Scott Owens: The Jackson Clinic, Jackson, TN, USA. Dr. Aftab Naz: Madera Family Medical Group, Madera, CA, USA. Dr. Juan Carillo: Buena Salud Pediatrics, San Jose, CA, USA. Dr. Patty J. Perry: Yuma Regional Outpatient Pediatrics Clinic, Yuma, AZ, USA. Dr. Lax Dedhia: Middletown Family Care Associates, Middletown, DE, USA. Dr. Michael A. Hassman: Comprehensive Clinical Research, Berlin, NJ, USA. Data Analysis: Data analysis was done by Yingjie Liu, James Gorsich and Rajesh Kasbekar. Randomization: Randomization was done by Michelle Secic (M.S Biostatistics) from Secic Consulting using a commercially available software application. Manuscript Preparation: The manuscript was prepared by Rajesh Kasbekar, Kristine Hendrikson, James Gorsich and Matt Baun in collaboration with Dr. Yingjie Liu, Dr Aftab Naz and Dr Lorenzo Marcos.

CONFLICT OF INTEREST

The study was implemented, and the data were collected by third-party independent investigators who had no financial interest in the products, technology or methodology discussed in this paper. The analysis and manuscript preparation were done by Kaz USA, Inc., a Helen of Troy Company in collaboration with the third-party investigators. The study protocol is available upon request.

ETHICAL APPROVAL

Approval for the study was obtained by a central Institutional Review Board (IRB), the Alpha IRB out of California in the US. Approvals for international sites were granted by the Human Subjects Protection Review Boards of the individual hospital prior to subject enrolment. This was in accordance with US code of federal regulations and EU clinical study requirements.

PATIENT CONSENT STATEMENT

Patient consent was obtained from adults as well as children between 7 and 17 years. Parent's or Guardian's consent was obtained for children below 7 years. This was in accordance with US code of federal regulations and EU clinical study requirements.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Lorenzo Marcos https://orcid.org/0000-0002-7332-0406

ENDNOTE

[1] STARD: Standards for reporting Diagnostic Accuracy Studies (File S1).

REFERENCES

Al-Eissa, Y. A., Al-Sanie, A. M., Al-Alola, S. A., Al-Shaalan, M. A., Ghazal, S. S., Al-Harbi, A. H., & Al-Wakeel, A. S. (2000). Parental perceptions of fever in children. *Annals of Saudi Medicine*, *20*(3-4), 202–205.

ASTM International (2016). *ASTM E 1965–98 (Reapproved 2016): Standard specification for infrared thermometers for intermittent determination of patient temperature*. American Society for Testing and Materials. Retrieved from https://reference.globalspec.com/standard/3862703/astm-e1965-98-2016

Avner, J. R. (2009). Acute fever. *Pediatrics in Review*, *30*(1), 5–13. https://doi.org/10.1542/pir.30-1-5

Baker, M. D., Bell, L. M., & Avner, J. R. (1993). Outpatient management without antibiotics of fever in selected infants. *New England Journal of Medicine*, *329*(20), 1437–1441. https://doi.org/10.1056/NEJM199311113292001

Baraff, L. J., Bass, J. W., Fleisher, G. R., Klein, J. O., McCracken, G. H. Jr, Powell, K. R., & Schriger, D. L. (1993). Practice guideline for the management of infants and children 0 to 36 months of age with fever without source. Agency for Health Care Policy and Research. *Annals of Emergency Medicine*, *22*(7), 1198–1210.

Barbi, E., Marzuillo, P., Neri, E., Naviglio, S., & Krauss, B. S. (2017). Fever in children: Pearls and pitfalls. *Children (Basel)*, *4*(9), 81. https://doi.org/10.3390/children4090081

Batra, P., Saha, A., & Faridi, M. M. (2012). Thermometry in children. *Journal of Emergencies, Trauma, and Shock*, *5*(3), 246–249. https://doi.org/10.4103/0974-2700.99699

Bettinelli, A., Provero, M. C., Cogliati, F., Villella, A., Marinoni, M., Saettini, F., Bianchetti, M. G., Nespoli, L., Galluzzo, C., & Lava, S. A.

(2013). Symptomatic fever management among 3 different groups of pediatricians in Northern Lombardy (Italy): Results of an explorative cross-sectional survey. *Italian Journal of Pediatrics*, *39*, 51. https://doi.org/10.1186/1824-7288-39-51

Black, R. E., Laxminarayan, R., Temmerman, M., & Walker, N. (2016). *Reproductive, maternal, newborn, and child health, Chapter 8 diagnosis and treatment of the febrile child* (pp. 137–161). World Bank. https://doi.org/10.1596/978-1-4648-0348-2_ch1

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, *1*(8476), 307–310. https://doi.org/10.1016/S0022-3476(82)80291-6

Canadian Paediatric Society (2015). *Fever and temperature taking. Information for parents from Canada's paediatricians - caring for kids, community paediatric society October 2015*. Retrieved from https://www.caringforkids.cps.ca/handouts/health-conditions-and-treatments/fever_and_temperature_taking

Chiu, C. H., Lin, T. Y., & Bullard, M. J. (1997). Identification of febrile neonates unlikely to have bacterial infections. *The Pediatric Infectious Disease Journal*, *16*(1), 59–63. https://doi.org/10.1097/00006454-199701000-00013

Crain, E. F., & Shelov, S. P. (1982). Febrile infants: Predictors of bacteremia. *Journal of Pediatrics*, *101*, 686–689. https://doi.org/10.1016/S0022-3476(82)80291-6

Demir, F., & Sekreter, O. (2012). Knowledge, attitudes and misconceptions of primary care physicians regarding fever in children: A cross sectional study. *Italian Journal of Pediatrics*, *38*, 40. https://doi.org/10.1186/1824-7288-38-40

El-Radhi, A. S. (2012). Fever management: Evidence vs current practice. *World Journal of Clinical Pediatrics*, *1*(4), 29–33. https://doi.org/10.5409/wjcp.v1.i4.29

European Committee for Standardization (2003). *Clinical thermometers - Part 5: Performance of infrared ear thermometers (With Maximum Device) Boutique AFNOR*, Saint-Denis.

Giuliano, K. K., Giuliano, A. J., Scott, S. S., MacLachlan, E., Pysznik, E., Elliot, S., & Woytowicz, D. (2000). Temperature measurement in critically ill adults: A comparison of tympanic and oral methods. *American Journal of Critical Care*, *9*(4), 254–261. https://doi.org/10.4037/ajcc2000.9.4.254

Greene, J. W., Hara, C., O'Connor, S., & Aftemeier, W. A. 3rd (1981). Management of febrile outpatient neonates. *Clinical Pediatrics*, *20*(6), 375–380. https://doi.org/10.1177/000992288102000601

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*(17–18), 2509–2518. https://doi.org/10.1111/jocn.12060

Hausfater, P., Zhao, Y., Defrenne, S., Bonnet, P., & Riou, B. (2008). Cutaneous infrared thermometry for detecting febrile patients. *Emerging Infectious Diseases*, *14*, 1255–1258. https://doi.org/10.3201/eid1408.080059

Herzog, L., & Phillips, S. G. (2011). Addressing concerns about fever. *Clinical Pediatrics*, *50*(5), 383–390. https://doi.org/10.1177/0009922810385929

International Organization for Standardization (2017). *ISO 80601-2-56:2017 (E) Medical electrical equipment: Particular requirements for basic safety and essential performance of clinical thermometers for body temperature measurement*. ISO. Retrieved from https://www.iso.org/obp/ui/#iso:std:iso:80601:-2-56:ed-1:v1:en

Kuppermann, N., Fleisher, G. R., & Jaffe, D. M. (1998). Predictors of occult pneumococcal bacteremia in young febrile children. *Annals of Emergency Medicine*, *31*, 679–687. https://doi.org/10.1016/S0196-0644(98)70225-2

Laupland, K. B., Shahpori, R., Kirkpatrick, A. W., Ross, T., Gregson, D. B., & Stelfox, H. T. (2008). Occurence and outcome of fever in critically ill adults. *Critical Care Medicine*, *36*, 1531–1535. https://doi.org/10.1097/CCM.0b013e318170efd3

Mangat, J., Standley, T., Prevost, A., Vasconcelos, J., & White, P. (2010). A comparison of technologies used for estimation of body temperature. *Physiological Measurement*, *31*(9), 1105–1118. https://doi.org/10.1088/0967-3334/31/9/003

Mogensen, C. B., Wittenhoff, L., Fruerhoj, G., & Hansen, S. (2018). Forehead or ear temperature measurement cannot replace rectal measurements, except for screening purposes. *BMC Pediatrics*, *18*(1), 15. https://doi.org/10.1186/s12887-018-0994-1

NICE National Institute for Health and Care Excellence (2019). *Overview: Fever in under 5s: Assessment and initial management: Guidance., November 2019*. Retrieved from https://www.nice.org.uk/guidance/ng143

Niehues, T. (2013). The febrile child: Diagnosis and treatment. *Deutsches Aerzteblatt Online*, *110*(45), 764–773. https://doi.org/10.3238/arztebl.2013.0764

Nijman, R. G., Vergouwe, Y., Thompson, M., Van Veen, M., Van Meurs, A. H., Van der Lei, J., Steyerberg, E. W., Moll, H. A., & Oostenbrink, R. (2013). Clinical prediction model to aid emergency doctors managing febrile children at risk of serious bacterial infections: Diagnostic study. *BMJ*, *346*, f1706. https://doi.org/10.1136/bmj.f1706

Nimah, M. M., Bashesh, 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*(1), 48–55. https://doi.org/10.1097/01.PCC.0000185476.35550.B2

Pade, K. H. (2019). Points and pearls: Evaluation and management of the febrile young infant in the emergency department. *Pediatric Emergency Medical Practice*, *16*(7), e1–e2.

Palladino, L., Woll, C., & Aronson, P. L. (2019). Evaluation and management of the febrile young infant in the emergency department. *Pediatric Emergency Medical Practice*, *16*(7), 1–24.

Pantell, R. H., Newman, T. B., Bernzweig, J., Bergman, D. A., Takayama, J. I., Segal, M., Finch, S. A., & Wasserman, R. C. (2004). Management and outcomes of care of fever in early infancy. *JAMA*, *291*(10), 1203–1212. https://doi.org/10.1001/jama.291.10.1203

Pecoraro, V., Petri, D., Costantino, G., Squizzato, A., Moja, L., Virgili, G., & Lecenteforte, E. (2020). The diagnostic accuracy of digital, infrared and mercury-in-glass thermometers in measuring body temperature: a systematic review and network meta-analysis. *Internal and Emergency Medicine*, 1–13. https://doi.org/10.1007/s11739-020-02556-0

Pepe, M. S. (2003). *The statistical evaluation of medical tests for classification and prediction*. Oxford University Press.

Pusic, M. V. (2007). Clinical management of fever in children younger than three years of age. *Paediatrics & Child Health*, *12*(6), 469–472. https://doi.org/10.1093/pch/12.6.469

Pušnik, I., & Miklavec, A. (2009). Dilemmas in measurement of human body temperature. *Instrumentation Science & Technology*, *37*(5), 516–530. https://doi.org/10.1080/10739140903149061

Schmitt, B. D. (1980). Fever phobia: Misconceptions of parents about fevers. *American Journal of Diseases of Children*, *134*(2), 176–181. https://doi.org/10.1001/archpedi.1980.02130140050015

Shaw, K. N., Gorelick, M., McGowan, K. L., Yakscoe, N. M., & Schwartz, J. S. (1998). Prevalence of urinary tract infection in febrile young children in the emergency department. *Pediatrics*, *102*(2), e16. https://doi.org/10.1542/peds.102.2.e16

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

Teach, S. J., & Fleisher, G. R. (1997). Duration of fever and its relationship to bacteremia in febrile outpatients three to 36 months old. The Occult Bacteremia Study Group. *Pediatric Emergency Care*, *13*(5), 317–319. https://doi.org/10.1097/00006565-199710000-00004

Van den Bruel, A., Thompson, M. J., Haj-Hassan, T., Stevens, R., Moll, H., Lakhanpaul, M., & Mant, D. (2011). Diagnostic value of laboratory tests in identifying serious infections in febrile children: Systematic review. *BMJ*, *342*, d3082. https://doi.org/10.1136/bmj.d3082

Waalen, J., & Buxbaum, J. N. (2011). Is older colder or colder older? The association of age with body temperature in 18,630 individuals. *The Journals of Gerontology Series A*, *66A*(5), 487–492. https://doi.org/10.1093/gerona/glr001

Wyckoff, A. S. (2009). Thermometer use 101. *American Academy of Pediatrics*. AAP News, November, *30*(11), 29. Retrieved from https://www.aappublications.org/content/30/11/29.2

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Kasbekar R, Naz A, Marcos L, et al. Threshold for defining fever varies with age, especially in children: A multi-site diagnostic accuracy study. *Nurs Open*. 2021;8:2705–2721. https://doi.org/10.1002/nop2.831