Research

Gail Hayward, Jan Y Verbakel, Fatene Abakar Ismail, George Edwards, Kay Wang, Susannah Fleming, Gea A Holtman, Margaret Glogowska, Elizabeth Morris, Kathryn Curtis and Ann van den Bruel

Non-contact infrared versus axillary and tympanic thermometers in children attending primary care:

a mixed-methods study of accuracy and acceptability

Abstract

Background
Guidelines recommend measuring temperature in children presenting with fever using electronic axillary or tympanic thermometers. Non-contact thermometry offers advantages, yet has not been tested against recommended methods in primary care.

Aim
To compare two different non-contact infrared thermometers (NCITs) to axillary and tympanic thermometers in children aged ≤5 years visiting their GP with an acute illness.

Design and setting
Method comparison study with nested qualitative component.

Method
Temperature measurements were taken with electronic axillary (Welch Allyn SureTemp®), electronic tympanic (Braun Thermoscan®), NCIT Thermofocus® 0800, and NCIT Firhealth Forehead. Parents rated acceptability and discomfort. Qualitative interviews explored parents' experiences of the thermometers.

Results
In total, 401 children were recruited (median age 1.6 years, 50.62% male). Mean difference between the Thermofocus NCIT and axillary thermometer was −0.14°C (95% confidence interval [CI] = −0.21 to −0.06°C); lower limit of agreement was −1.57°C (95% CI = −1.69 to −1.44°C) and upper limit 1.29°C (95% CI = 1.16 to 1.42°C). A second NCIT (Firhealth) had similar levels of agreement; however, the limits of agreement between tympanic and axillary thermometers were also wide. Parents expressed a preference for the practicality and comfort of NCITs, and were mostly negative about their child's experience of axillary thermometers. But there was willingness to adopt whichever device was medically recommended.

Conclusion
In a primary care paediatric population, temperature measurements with NCITs varied by >1°C compared with axillary and tympanic approaches. But there was also poor agreement between tympanic and axillary thermometers. Since clinical guidelines often rely on specific fever thresholds, clinicians should interpret peripheral thermometer readings with caution and in the context of a holistic assessment of the child.

Keywords
acute disease; child; fever; primary health care; thermometers.

INTRODUCTION

Acute infection in children is one of the most common problems in general practice and is associated with considerable burden on NHS resources. Nearly 40% of parents with children aged 6–17 months consult a healthcare professional when their child has a high temperature.[1] In the UK, acute infections result in 4 consultations per person-year in children aged <1 year, and 1.3 consultations per person-year in children aged 1–15 years.[2] Febrile illness accounts for 20% of all visits to the paediatric emergency department.[3]

In children aged ≥4 weeks presenting with fever symptoms, guidelines recommend that the measurement of temperature should be taken with electronic axillary thermometers or with tympanic thermometers.[4] However, axillary thermometers require healthcare professionals to undress the child and hold the thermometer in the axilla for >15 seconds.[5] Tympanic thermometers are easier to use, but may be inaccurate because of the presence of ear wax or insufficient straightening of the ear canal,[6] and both types of device require disposable covers to avoid cross-infection. Non-contact infrared thermometers (NCITs) convert measurements of the intensity of infrared radiation emitted by the body into temperature readings. The non-contact approach offers potential advantages, including reduced child discomfort and distress, rapid readings, measurement without interrupting sleep, minimal risk of cross-infection, and no requirement for disposable covers.[7]

Reports of agreement between NCITs and conventional thermometers have been variable, with larger mean differences reported between NCITs and tympanic thermometers than with NCITs and mercury-in-glass axillary thermometers.[8,9] Comparisons between electronic and mercury rectal thermometry have also yielded variable results,[10,11] and performance varies between devices.[12] Finally, although NCITs are mostly reported to have high sensitivity and specificity in detecting a fever of ≥38°C measured with conventional thermometry,[8–10,13] in two studies sensitivity was estimated as only 27%[11] and 12%.[12]

In addition to the lack of clear conclusions from existing studies, there is a lack of generalisability of this data to primary

G Hayward (ORCID: 0000-0003-0852-627X), DPhil, MRCGP, associate professor; **G Edwards** (ORCID: 0000-0002-0048-4278), MSc, research assistant; **K Wang** (ORCID: 0000-0002-7195-1730), DPhil, MRCGP, National Institute for Health Research (NIHR) postdoctoral fellow; **S Fleming** (ORCID: 0000-0001-7205-2051), DPhil, quantitative researcher; **M Glogowska** (ORCID: 0000-0001-8029-1052), PhD, senior qualitative researcher; **E Morris** (ORCID: 0000-0002-9913-6041), MRCGP, NIHR in-practice fellow, Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK. **JY Verbakel** (ORCID: 0000-0002-7166-7211), PhD, MD, honorary clinical researcher, assistant professor; **A van den Bruel**, PhD, MD, associate professor, Academic Centre for Primary Care, Department of Public Health and Primary Care, KU Leuven, Leuven, Belgium. **F Abakar Ismail** (ORCID: 0000-0002-1463-7980), PhD, research associate, Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow Royal Infirmary, Glasgow, UK. **GA Holtman** (ORCID: 0000-0001-6579-767X), PhD, academic visitor, postdoctoral researcher, Department of General Practice and Elderly Care Medicine, University Medical Centre Groningen, University of Groningen, Groningen, the Netherlands. **K Curtis** (ORCID: 0000-0001-5367-9182), BA, trial manager, Bristol Medical School: Population Health Sciences, University of Bristol, Bristol, UK.

Address for correspondence
Gail Hayward, Nuffield Department of Primary Care Health Sciences, Radcliffe Observatory Quarter, Woodstock Road, Oxford OX2 6GG, UK.

Email: gail.hayward@phc.ox.ac.uk

How this fits in

Non-contact infrared thermometers (NCITs) allow clinicians to measure temperature from the forehead without physical contact, reducing discomfort and distress, and the need for disposable covers. However, NCITs have not been tested against the current recommended methods in primary care. Two different NCITs were compared with electronic axillary and tympanic thermometers in children aged ≤5 years visiting their GP with an acute illness. Analysis of the results suggests that, in 95% of cases, the difference between NCITs and electronic axillary or tympanic thermometers was up to 1°C. Since the results show poor agreement between axillary and tympanic thermometer readings, the precision of each thermometer type is unknown. Clinicians should be aware of the variability in peripheral thermometer readings when assessing febrile children according to clinical guidelines.

care settings. Most previous studies were conducted in paediatric inpatient populations,[9,11,12] or mixed hospital ambulatory care and ward settings.[10,14] Furthermore, NCITs have been mainly compared with temperature measurement approaches that are not currently recommended for use in children, including rectal measurements[10–12,15] and using mercury-in-glass axillary thermometers.[8]

Understanding the performance of NCITs, compared with the currently recommended tympanic and electronic axillary thermometers in a primary care paediatric population, could support introduction of this potentially beneficial technology into routine practice. This mixed-methods study evaluated the agreement between two NCIT models with electronic axillary and tympanic thermometers in children who present with acute illness in primary care, and assessed their acceptability to parents and children.

METHOD

This was a cross-sectional method agreement study with a nested qualitative study.

Method comparison

Children aged between 0–5 years with an acute illness and symptoms first presenting within 14 days visiting a GP practice (nine sites) or an out-of-hours (OOH) service (one site) in Oxfordshire, UK, were eligible for inclusion. Children for whom acute trauma was the main reason for presentation; who were clinically unstable; had already participated in the study; or where parents were unable to understand the study materials (written in English) were excluded from the study.

Parents and their children were approached consecutively in the OOH or surgery waiting room by a study researcher between April 2017 and August 2018. Parents were provided with an information sheet and gave verbal consent to their child's involvement, with the option to consent to further contact by telephone for involvement in the qualitative substudy, and the option to withdraw consent subsequently by email or telephone. Temperature measurements were conducted prior to or after the child's GP appointment. Demographic information and the history of fever were recorded.

Four thermometers were compared: electronic axillary (Welch Allyn SureTemp®), electronic tympanic (Braun Thermoscan®), NCIT Thermofocus® 0800, and NCIT Firhealth Forehead Thermometer. The Thermofocus NCIT was included as it had been most extensively evaluated in other settings.[8,10,12,15,16] The Firhealth device was included as an example of a cheaper NCIT.

Measurements were performed consecutively in the shortest time frame possible, and no medication or drinks were administered between measurements. The order in which the thermometers were used was randomised prior to the clinical stage of the study for each participant using a random number generator (https://www.random.org). Tympanic measurements were not taken in babies aged <4 weeks, in line with UK guidance.[4] Once the four primary measurements were complete, a second measurement was taken with the Thermofocus and Firhealth thermometers to evaluate their reproducibility. Failed measurements were recorded due to lack of cooperation of the child after three attempts, mechanical issues (operational or technological failure), and clinically implausible readings (based on researcher's assessment).

The children's reaction to the different measurements was rated by parents using the Patient Discomfort Scale.[17] Children aged 4 and 5 years additionally completed the Wong–Baker Faces® Pain Rating Scale.[18] Parents scored the acceptability of each thermometer on a 10 cm Visual Analogue Scale before they were informed of the temperature measurements.

Sample size

The sample size calculation was based on the desired accuracy of the limits of agreement on the Bland–Altman plots for the primary outcome.[19] The limits of agreement are

defined as the mean difference (bias) ±1.96 standard deviations.[20] If the differences are normally distributed, approximately 95% of them are be expected to lie within the limits of agreement, so they are a useful estimate of the typical range of differences between the two measurements. Assuming an accuracy of ±0.075°C would be desired for the 95% confidence intervals (CIs) of the limits of agreement, and the standard deviation of the agreement between temperatures measured by NCIT and electronic axillary thermometers (based on previous studies)[8–10,13,15] would be 0.5°C, a minimum sample size of 533 participants would have been required. The sample size calculation was revised on 14 May 2018. Based on the data already available, the standard deviation for primary analysis (Thermofocus NCIT versus electronic axillary thermometer) was 0.65°C. Using this standard deviation, the original sample size of 533 children would give the authors a 0.10°C accuracy for each limit of agreement (rather than 0.075 as anticipated). A reduced sample size of 400 children would give the authors 0.11°C accuracy. Considering thermometers only measure temperatures <0.1°C, this would be sufficient as the rounding would make the two estimates equivalent. Secondary outcomes of agreement between the other thermometer types have been estimated with the same precision.

Analysis

Statistical methods focus on the agreement between thermometers, the accuracy of detecting fever, and failure rates. All children contributed data to each analysis, when available.

The primary outcome was the agreement between the Thermofocus NCIT and the electronic axillary thermometer. Analyses of agreement were conducted based on Bland–Altman plots,[19] which provided an indication of bias and 95% limits of agreement between the measurements. Ninety-five per cent CIs around these estimates have been calculated.

Diagnostic accuracy for detecting fever (temperature ≥38°C measured by the electronic axillary thermometer) was analysed by calculating sensitivity, specificity, predictive values, and likelihood ratios, with 95% CIs. Failure rates are reported as proportions.

The scores on the Visual Analogue and Patient Discomfort Scales have been analysed using non-parametric techniques resulting in median acceptability (and interquartile ranges [IQRs]) for each thermometer.

Qualitative data collection and analysis

Parents who consented to contact were purposively sampled to achieve maximum variation in sex of parent, age of parent, age of child, ethnicity, and number of siblings. Recruitment continued until the research team agreed data saturation had been achieved and sufficient explanation for the categories generated was reached.

Interviews were semi-structured and conducted by telephone ($n = 20$) or face-to-face ($n = 1$) using a flexible topic guide developed by the research team and patient and public involvement (PPI) panel, which evolved in response to emerging themes. The topic guide explored parents' experiences with the different thermometers, their thoughts about future use of thermometers, and wider exploratory questions about motivations for and experience of temperature measurement and fever in children. A detailed example of the topic guide is presented in Supplementary Box S1.

All participants gave written or recorded verbal informed consent prior to the interview. Interviews were conducted separately by two researchers trained in qualitative methodology (a female clinical researcher and salaried GP, and a female research assistant), and were audio-

Table 1. Participant characteristics, *N* = 401

Age, years, median (IQR)	1.6 (0.79–3.38)
Sex, male, *n* (%)	203 (50.62)
Ethnicity, *n* (%)	
White British	280 (69.83)
White other	38 (9.48)
Mixed	27 (6.73)
Pakistani	21 (5.24)
Other Asian	11 (2.74)
African	8 (2.00)
Indian	5 (1.25)
Chinese	5 (1.25)
Bangladeshi	4 (1.00)
Caribbean	1 (0.25
Black British	1 (0.25)
Mother's age, years, median (IQR)	32 (29–36)
Number of siblings, median (IQR)	1 (0–1)
Parent believed child to be febrile at point of assessment, *n* (%)[a]	119 (29.75)
Fever medication in past 6 hours, *n* (%)[a]	134 (33.50)
Parent report of fever duration, days, median (IQR)	1.5 (0.5–3)
Illness duration, days, median (IQR)	3 (2–7)
Recruitment site, *n* (%)	
OOH primary care	34 (8.48)
In-hours primary care	367 (91.52)

[a]N = 400 as there was missing data point for one participant for these questions. IQR = interquartile range.
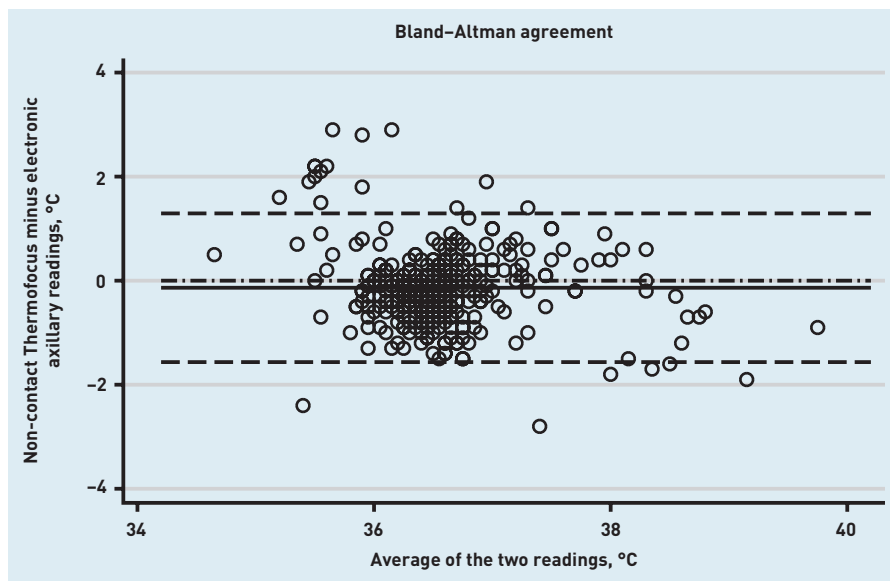
OOH = out-of-hours.

Bland–Altman agreement

Figure 1. Bland–Altman plot for agreement between the Thermofocus NCIT thermometer and the electronic axillary thermometer.
Solid line: mean difference between the two methods; dashed lines: upper and lower limits of agreement; dash-dotted line: line of no difference.

recorded and transcribed. Consistency was ensured by regular discussion and review of transcripts and topic guides by the team.

Data analysis followed a thematic approach, with the assistance of NVivo (version 11). This included familiarisation with the data, open coding, and subsequent inductive reasoning to identify salient categories and relationships between emerging themes derived from the data. Data and codes were then checked by two researchers. The codes and themes were developed and interpreted in discussion with the wider research team.

Table 2. Method agreement results between the Thermofocus NCIT, Firhealth NCIT, axillary, and tympanic thermometers

	Mean difference and 95% CI (°C)	Lower limit of agreement and 95% CI (°C)	Upper limit of agreement and 95% CI (°C)
Method comparison			
Thermofocus NCIT minus electronic axillary, *n* = 371	–0.14 (–0.21 to –0.06)	–1.57 (–1.69 to –1.44)	1.29 (1.16 to 1.42)
Firhealth NCIT minus electronic axillary, *n* = 374	–0.16 (–0.23 to –0.09)	–1.54 (–1.66 to –1.41)	1.22 (1.10 to 1.34)
Thermofocus NCIT minus tympanic, *n* = 386	–0.10 (–0.17 to –0.03)	–1.55 (–1.68 to –1.42)	1.35 (1.22 to 1.48)
Firhealth NCIT minus tympanic, *n* = 389	–0.10 (–0.17 to –0.03)	–1.47 (–1.59 to –1.35)	1.28 (1.16 to 1.40)
Thermofocus NCIT minus Firhealth NCIT, *n* = 395	0.00 (–0.04 to 0.05)	–0.90 (–0.98 to –0.82)	0.91 (0.83 to 0.99)
Electronic axillary minus tympanic, *n* = 367	0.06 (–0.02 to 0.14)	–1.49 (–1.63 to –1.34)	1.61 (1.47 to 1.75)
Reproducibility			
Thermofocus NCIT 1st minus 2nd reading, *n* = 395	–0.04 (–0.07 to –0.01)	–0.56 (–0.60 to –0.51)	0.47 (0.43 to 0.52)
Firhealth NCIT 1st minus 2nd reading, *n* = 397	0.01 (–0.02 to 0.04)	–0.60 (–0.65 to –0.54)	0.61 (0.56 to 0.67)

Patient and public involvement

The PPI panel for the National Institute for Health Research Community Healthcare MedTech and In Vitro Diagnostics Cooperative, which includes two mothers of children aged <5 years, supported this project from inception to dissemination. The PPI panel have provided feedback on the study design, suggesting the inclusion of reproducibility as a secondary outcome and endorsing the inclusion of a cheaper NCIT as a more affordable option for home use. They also provided feedback on study materials, particularly the patient information leaflets and acceptability ratings by parents and their children, commented on the topic guide for the qualitative interviews with the parents, and provided feedback on the sample size revision. The PPI panel provided advice on acceptable limits of agreement between thermometry devices, and discussed the emerging themes on the acceptability of the thermometer types from the qualitative interviews. The panel also provided advice regarding dissemination of the findings to a lay audience.

RESULTS

In total, 401 children were recruited with a median age of 1.6 years (IQR 0.79 to 3.38 years); 203 (50.62%) were male. Five children were <4 weeks old. Most of the children were of white British ethnicity (69.83%). At the time of inclusion, 29.68% of the children were believed by their parents to be feverish at the time of inclusion. Participant characteristics are listed in Table 1.

The Bland–Altman plot for the primary outcome, which is the agreement between the Thermofocus NCIT (first measurement) and the electronic axillary thermometer, is presented in Figure 1. The mean difference between the two methods was –0.14°C (95% CI = –0.21 to –0.06°C), with the lower limit of agreement being –1.57°C (95% CI = –1.69 to –1.44°C) and the upper limit being 1.29°C (95% CI = 1.16 to 1.42°C). This means that in 95 out of 100 cases the difference between the NCIT and electronic axillary thermometer would be between 1.57°C lower and 1.29°C higher than the average of the NCIT and electronic axillary measurements of temperature.

The mean difference between the Firhealth NCIT and the electronic axillary thermometer was –0.16°C (95% CI = –0.23 to –0.09°C); the lower limit of agreement was –1.54°C (95% CI = –1.66 to –1.41°C) and the upper limit was 1.22°C (95% CI = 1.10 to 1.34°C). Full method comparison results, including NCIT versus tympanic and repeated measurements with both
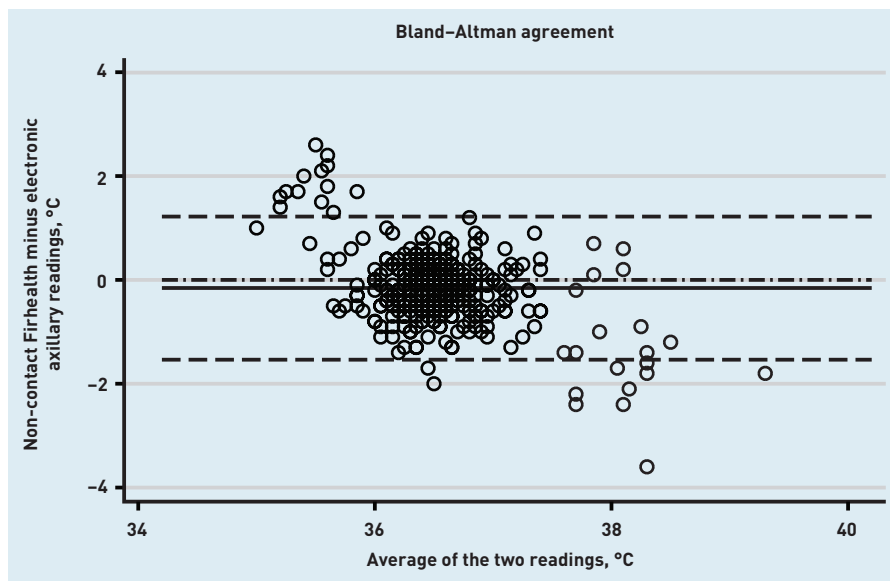
Bland–Altman agreement

Figure 2. Bland–Altman plot for agreement between the Firhealth NCIT thermometer and the electronic axillary thermometer.
Solid line: mean difference between the two methods; dashed lines: upper and lower limits of agreement; dash-dotted line: line of no difference.

NCITs are summarised in Table 2 (see Supplementary Figure S1 for accompanying plots). Of note, there were wide limits of agreement when the tympanic and electronic axillary methods were compared. The mean difference was 0.06°C (95% CI = −0.02 to 0.14°C); the lower limit of agreement was −1.49°C (95% CI = −1.63 to −1.34°C) and the upper limit of agreement was 1.61°C (95% CI = 1.47 to 1.75°C). The reproducibility of the NCITs was reasonable; for the Thermofocus, the mean difference was −0.04°C (95% CI = −0.07 to −0.01°C); the lower limit of agreement was −0.56°C (95% CI = −0.60 to −0.51°C) and the upper limit of agreement was 0.47°C (95% CI = 0.43 to 0.52°C), meaning that the two measurements varied by less than 0.5°C in most cases. Firhealth reproducibility was similarly acceptable. Although there was some indication in the Bland–Altman plots of temperature-related relationships in the difference between measurements made by NCITs and axillary thermometry (Figures 1 and 2), no further investigations were performed as they are dominated by a small number of measurements at extremes of the physiological temperature range, where it is expected that thermometers for physiological use will be less accurate.

The accuracy of NCITs for diagnosing fever was calculated, defined as a temperature of ≥38°C by electronic axillary measurement. Based on the axillary thermometer, there were 41 children with a body temperature of ≥38°C, which was a prevalence of fever of 10.22%. The proportion of children detected as febrile by the Thermofocus NCIT out of all children found to be febrile using the electronic axillary thermometer (that is,

the sensitivity for this alternate imperfect reference standard) was 29.3% (95% CI = 16.1 to 45.5%), and the proportion found not to be febrile out of all the children classified as not febrile by electronic axillary measurement (specificity) was 96.9% (95% CI = 94.6 to 98.5%). For the Firhealth NCIT, sensitivity was 4.9% (95% CI = 0.6 to 16.5%) and specificity 98.9% (95% CI = 97.2 to 99.7%) (see Table 3 for further details).

The number of attempts for each thermometer that were required to obtain a valid reading and the technical failures are detailed in Table 4.

Thermometer acceptability
Children (*n* = 69) aged 4 or 5 years completed the Wong–Baker Faces® Pain Rating Scale. The median score and IQR was 0 (0 to 0) for both NCITs, and 0 (IQR 0 to 2) for the electronic axillary and tympanic thermometers.

Overall, most children were rated by their parents as relaxed during temperature measurements with each thermometer. The median Patient Discomfort Scale (completed by 306 parents) for each thermometer was 2, although the IQR was slightly larger for the electronic axillary thermometer than for the other thermometers: Thermofocus median 2 (IQR 2 to 2), Firhealth median 2 (IQR 2 to 2), electronic axillary median 2 (IQR 2 to 3), and tympanic median 2 (IQR 2 to 2).

Parental acceptability as assessed by a 10 cm Visual Analogue Scale by 398 parents was highest for the Firhealth NCIT (median 9.0 cm; IQR 7.6 to 9.5 cm), followed by the Thermofocus NCIT (median 8.5 cm; IQR 6.9 to 9.4 cm), the tympanic thermometer (median 7.6 cm; IQR 5.5 to 9.0 cm), and lastly the electronic axillary thermometer (median 5.0 cm; IQR 2.1 to 7.6).

Qualitative interviews
The characteristics of 21 parents who participated in the interviews are described in Supplementary Table S1. Themes relating to device attributes — what parents look for and feel they need from thermometers — are presented here. Key themes were convenience and practicality, comfort, cost, safety, and endorsement. Further qualitative data and analysis will be reported separately in a future publication.

Electronic axillary thermometer
Parents' experiences with the axillary thermometers were described in almost all cases using negative language, with over half describing their child as appearing 'uncomfortable' during its use. For some, the appearance of the device contributed to the child's negative experience:

Table 3. Diagnostic accuracy for fever defined as ≥38°C by the electronic axillary thermometer

	Thermofocus	Firhealth	Tympanic
Sensitivity, % (95% CI)	29.3 (16.1 to 45.5)	4.9 (0.6 to 16.5)	29.3 (16.1 to 45.5)
Specificity, % (95% CI)	96.9 (94.6 to 98.5)	98.9 (97.2 to 99.7)	93.6 (90.6 to 95.9)
Positive predictive value, % (95% CI)	52.2 (30.6 to 73.2)	33.3 (4.3 to 77.7)	34.3 (19.1 to 52.2)
Negative predictive value, % (95% CI)	92.3 (89.2 to 94.8)	90.1 (86.8 to 92.9)	92.1 (88.8 to 94.6)
Likelihood ratio + (95% CI)	9.58 (4.52 to 20.31)	4.39 (0.83 to 23.24)	4.58 (2.47 to 8.50)
Likelihood ratio – (95% CI)	0.73 (0.60 to 0.89)	0.96 (0.90 to 1.03)	0.76 (0.62 to 0.92)
Absolute numbers (TP, FP, FN, TN)	12 11 29 349	2 4 29 356	12 23 29 337

FN = false negative. FP = false positive. TN = true negative. TP = true positive.

'He freaked out at that one … Because he thought it was a needle … so he thought he was going to have an injection.' (Interview participant [I]8, Mother, aged 31–40 years, two children)

Parents raised concerns about the practicality of this device; however, if it was the medically recommended thermometer, they would be willing to persevere:

'From a practical and user point of it, you know, it just seemed impractical, but if it's the best then so be it, you know, there was no harm.' (I1, Father, two children)

Tympanic thermometer
Parents' views on tympanic thermometers were more neutral than axillary thermometers, with several parents using them at home. They commonly expressed concern that thermometer performance might be affected by their failure to operate it correctly:

Table 4. Number of attempts needed to achieve first measurement and technical failures for each thermometer

	Thermofocus NCIT, *n* (%)	Firhealth NCIT, *n* (%)	Electronic axillary, *n* (%)	Tympanic, *n* (%)
One attempt required	382 (95.3)	390 (97.3)	363 (90.5)	364 (90.8)
Two attempts required	10 (2.5)	8 (2.0)	11 (2.7)	15 (3.7)
Three attempts required	4 (1.0)	1 (0.2)	2 (0.5)	9 (2.2)
No reading				
Technical error (thermometer not activating)	3 (0.7)	1 (0.2)	0 (0.0)	0 (0.0)
Technical error (other)	1 (0.2)	0 (0.0)	7 (1.7)	3 (0.7)
Lack of cooperation of the child	1 (0.2)	0 (0.0)	16 (4.0)	5 (1.2)
Reason not specified	0 (0.0)	1 (0.2)	2 (0.5)	0 (0.0)
Thermometer unsuitable	0 (0.0)	0 (0.0)	0 (0.0)	5 (1.2)

NCIT = non-contact infrared thermometer.

'I'm sure if I got it, if I was trained I might know better but I'm never sure if I've put it far enough into her ear or not far enough in or, I mean I'd always had a normal reading from it but I'm not confident that I wouldn't miss something by not being able to use it perfectly.' (I17, Mother, aged 31–40 years, one child)

Non-contact infrared thermometers
Parents highlighted the practicality and convenience of NCITs. One of the NCIT devices provided feedback to parents about the correct distance from their child's forehead they should hold the device to ensure an accurate reading. The feedback regarding this placement was reported to be helpful to ensure 'correct' usage and interpretation:

'It was as if nothing was happening he was just sitting on my lap, he was perfectly calm, so they were, I was quite impressed with those actually, I didn't know they existed.' (I10, Father, one child, with history of fever this episode)

'With the blue one there was a light, I thought that was quite helpful so you knew exactly the distance you needed to be … I guess that also helps with the accuracy of reading because you know you're taking it the right distance …' (I15, Mother, aged 31–40 years, one child)

Features of an ideal thermometer
Parents described five key attributes of thermometers that would influence their likelihood of using them: convenience of use, comfort for their child, and the cost, safety, and endorsement of the device. These themes were reflected throughout the descriptions of their child's experiences with the different devices.

Parents also displayed interest in the cost of devices for themselves and for the NHS:

'The non-contact … it's a lot easier, there's less parts to exchange on it, whereas with the underarm you've got to wipe it, with the ear one you've got to replace the caps so, you know, it all comes down to cost. Every patient comes in and you have to replace one cap for each patient you do an ear thermometer, you know, it all adds up doesn't it.' (I3, Father, one child)

However, some parents required reassurance regarding the reliability and safety of novel, or less familiar, technology:

'I think it's mainly because you're kind of used to ... like that physical contact for 1 or 2 minutes, so if you kind of use anything that is very different from the usual way of doing things then it always comes to your mind is this really reliable ...' (I14, Father, aged 31–40 years, two children, with history of fever this episode)

'I didn't mind doing it, I did ask a couple of times if it was safe to use all of them ... it was just because your colleague said it was infrared and they're new they're trialling them so I thought oh I'll ask the question it is on a child ...' (I4, Mother, aged 31–40 years, two children)

DISCUSSION

Summary

In this method comparison study, although the mean difference between NCIT and other thermometer measurements was only moderate, the upper and lower limits of agreement were >1°C. This has the potential to adversely affect clinical decision making, given that guidance relates to thresholds of temperature.[4] This also exceeds the limits of agreement of ±0.5°C, which are commonly believed to be acceptable.[21] The proportion classified as febrile out of all those with an electronic axillary measurement of ≥38°C was low for both NCITs. The producibility of measurements with NCITs was good. Fewer attempts were required and fewer failures were reported for NCITs, and performance was similar for the cheaper and more expensive brands tested.

The majority of parents rated all devices as acceptable, although parental satisfaction and child discomfort appear to be better for the NCITs. The axillary thermometer was least liked by parents and their children, and resulted in the most failed readings. When interviewed, parents expressed a preference for the practicality and comfort afforded by NCITs, and were predominantly negative about the user experience of axillary thermometers, which were felt to be more intrusive and have the potential to cause distress or discomfort, particularly if their child was unwell. Parents considered the convenience of use, the comfort of their child, cost, safety, and endorsement when evaluating thermometers; however, there was a willingness to adopt whichever device was medically recommended.

Strengths and limitations

To the authors' knowledge, this was the first method comparison study of peripheral thermometers to be conducted in children attending primary care with acute illness. In total, 401 children were recruited, making this one of the larger studies to address this question. However, there are some limitations to the study. First, uncertainty exists over the accuracy of the recommended electronic axillary and tympanic methods of gauging temperature. This study found wide limits of agreement between tympanic and axillary thermometers in this population. A recent systematic review[22] found 18 studies comparing tympanic thermometers and 19 studies comparing axillary thermometers to core temperature measurement (which was rectal measurement in the majority of cases) in children, with no studies conducted in primary care. The pooled estimates of mean difference was −0.43°C (95% limits of agreement −1.40 to 0.55°C) for axillary and −0.15°C (95% limits of agreement −0.67 to 0.37°C) for tympanic, although heterogeneity was very high and findings varied when different brands were considered separately. The authors concluded that neither method met acceptable limits of agreement compared with core temperature measurements. Therefore, it is impossible to be certain which of the thermometers was closest to the core temperature; NCITs could perform better than established methods.

Furthermore, the diagnostic accuracy of axillary and tympanic thermometers for fever using core body temperature measurement as the reference standard is variable. A systematic review found five paediatric studies reporting sensitivities ranging from 14% to 63% for electronic axillary measurement, and 13 paediatric studies reported sensitivities ranging from 0.23 to 0.87 for tympanic measurement of fever of ≥38°C. This complicates the interpretation of the estimates of the sensitivity and specificity of NCITs.

Second, as just over 10% of the participants were febrile on study entry, the CIs around the estimate of sensitivity are wide. The research team approached consecutive children arriving at the GP, but the parents who agreed to participate may have been those with children they felt to be less unwell, which may have resulted in under-recruitment of febrile children. However, over 30% of recruited children had received medication for fever in the previous 6 hours, and, by limiting the recruited population to children with recent onset of illness, those most likely to require temperature measurement as part of their assessment were included. Finally, the temperature measurements were performed by study personnel who inevitably developed expertise in using the equipment, which means that failure rates and accuracy may differ from

Discuss this article
Contribute and read comments about this article: **bjgp.org/letters**

what would be expected in practice staff or parents. However, all thermometers were simple to use and frequent use in primary care settings would ensure similar expertise developed.

Comparison with existing literature
There have been no systematic reviews evaluating the accuracy of NCITs; however, the poor performance of electronic axillary and tympanic thermometers compared with core body temperature measurements has been highlighted in a number of reviews.[6,22–25] The mean difference of –0.14°C and –0.16°C between NCITs and electronic axillary thermometers found in this study was within the range demonstrated by other studies comparing NCITs with electronic axillary,[26] tympanic,[9] electronic,[11,12] and mercury rectal thermometers.[10] The only study reporting comparable data for the Thermofocus brand compared this NCIT with a mercury in-glass thermometer used in the axilla in ambulatory paediatrics, in which a greater agreement was found, with an overall mean difference of 0.07°C and limits of agreement of –0.62°C (–0.67 to –0.47°C) and 0.76°C (0.61 to 0.91°C).[8] However, they used the average of two axillary measurements and three NCIT measurements in each child, which could have contributed to the improved performance. Both NCITs had very low sensitivity for fever as detected by axillary thermometry. Two studies, using Thermoflash® and Beurer NCITs, have demonstrated low sensitivity for fever as detected by rectal thermometry.[11,12] However, five studies,[9–11] two using the Thermofocus,[8,12] have found much higher sensitivity for fever.

While parents' knowledge and beliefs about fever and temperature management in their children have previously been described,[26,27] there was little description of the impact or experience of different thermometers. Parental concerns regarding practicality and comfort were highlighted in a study exploring attitudes to rectal thermometers.[28]

Implications for research and practice
In children aged ≤5 years, temperature measured by NCITs can vary by >1°C from measurements made by axillary and tympanic thermometers, which is a potentially clinically significant variability. Given the uncertainty over the accuracy of electronic axillary and tympanic thermometers for core body temperature, it was hard to draw firm conclusions about the likely impact on practice if NCITs were introduced as standard care.

If axillary thermometry is assumed to be an accurate reference standard, then the moderate agreement between NCITs and axillary, and the finding that a different population of children would be classified as febrile by NCITs compared with electronic axillary thermometers, means that decisions about the best pathway of care for a child could vary depending on which thermometer is used.[4] This in turn suggests that primary care clinicians should be cautious in using this technology.

However, this study also found wide limits of agreement between the electronic axillary and tympanic thermometers, both of which are advocated in guidance, and existing studies cast doubt over the accuracy of either method compared with core body temperature measurement.[22] Furthermore, parents were more positive about the benefits of NCITs than other types of devices in both rating scales and qualitative interviews, and there was good agreement between repeated measurements in both NCITs. A high-quality study comparing NCITs, axillary, and tympanic measurements to core body temperature measurements could help clarify this, but the invasive nature of core body temperature measurement will make this a challenge to conduct in a primary care population.[29]

Therefore, clinicians need to be cautious about the accuracy of any peripheral thermometry approach and ensure that the management decisions they make are using this data as part of a holistic assessment. There is clear potential for technological innovation in this field to develop more accurate methods of peripheral thermometry to support clinical decision making.

REFERENCES

1. Hay AD, Heron J, Ness A. The prevalence of symptoms and consultations in pre-school children in the Avon Longitudinal Study of Parents and Children (ALSPAC): a prospective cohort study. *Fam Pract* 2005; **22(4):** 367–374.

2. Fleming DM, Smith GE, Charlton JR, *et al.* Impact of infections on primary care — greater than expected. *Commun Dis Public Health* 2002; **5(1):** 7–12.

3. Armon K, Stephenson T, Gabriel V, *et al.* Determining the common medical presenting problems to an accident and emergency department. *Arch Dis Child* 2001; **84(5):** 390–392.

4. National Institute for Health and Clinical Excellence. *Fever in under 5s: assessment and initial management. NG143.* London: NICE, 2019. https://www.nice.org.uk/guidance/NG143 (accessed 3 Mar 2020).

5. NHS. *How to take your baby's temperature: your pregnancy and baby guide.* 2020. https://www.nhs.uk/conditions/pregnancy-and-baby/how-to-take-your-babys-temperature (accessed 3 Mar 2020).

6. Craig JV, Lancaster GA, Taylor S, *et al.* Infrared ear thermometry compared with rectal thermometry in children: a systematic review. *Lancet* 2002; **360(9333):** 603–609.

7. Wang K, Gill P, Wolstenholme J, *et al.* Non-contact infrared thermometers for measuring temperature in children: primary care diagnostic technology update. *Br J Gen Pract* 2014; DOI: https://doi.org/10.3399/bjgp14X682045.

8. Chiappini E, Sollai S, Longhi R, *et al.* Performance of non-contact infrared thermometer for detecting febrile children in hospital and ambulatory settings. *J Clin Nurs* 2011; **20(9–10):** 1311–1318.

9. Ng DK, Chan CH, Lee RS, Leung LC. Non-contact infrared thermometry temperature measurement for screening fever in children. *Ann Trop Paediatr* 2005; **25(4):** 267–275.

10. Teran CG, Torrez-Llanos J, Teran-Miranda TE, *et al.* Clinical accuracy of a non-contact infrared skin thermometer in paediatric practice. *Child Care Health Dev* 2012; **38(4):** 471–476.

11. Allegaert K, Casteels K, van Gorp I, Bogaert G. Tympanic, infrared skin, and temporal artery scan thermometers compared with rectal measurement in children: a real-life assessment. *Curr Ther Res Clin Exp* 2014; **76:** 34–38.

12. Paes BF, Vermeulen K, Brohet RM, *et al.* Accuracy of tympanic and infrared skin thermometers in children. *Arch Dis Child* 2010; **95(12):** 974–978.

13. Selent MU, Molinari NM, Baxter A, *et al.* Mass screening for fever in children: a comparison of 3 infrared thermal detection systems. *Pediatr Emerg Care* 2013; **29(3):** 305–313.

14. Osio CE, Carnelli V. Comparative study of body temperature measured with a non-contact infrared thermometer versus conventional devices. The first Italian study on 90 pediatric patients. *Minerva Pediatr* 2007; **59(4):** 327–336.

15. Fortuna EL, Carney MM, Macy M, *et al.* Accuracy of non-contact infrared thermometry versus rectal thermometry in young children evaluated in the emergency department for fever. *J Emerg Nurs* 2010; **36(2):** 101–104.

16. Sollai S, Dani C, Berti E, *et al.* Performance of a non-contact infrared thermometer in healthy newborns. *BMJ Open* 2016; **6(3):** e008695.

17. Greenes DS, Fleisher GR. Accuracy of a noninvasive temporal artery thermometer for use in infants. *Arch Pediatr Adolesc Med* 2001; **155(3):** 376–381.

18. Wong DL, Baker CM. Pain in children: comparison of assessment scales. *Pediatr Nurs* 1988; **14(1):** 9–17.

19. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; **1(8476):** 307–310.

20. Bland M. *How can I decide the sample size for a study of agreement between two methods of measurement?* 2004. https://www-users.york.ac.uk/~mb55/meas/sizemeth.htm (accessed 3 Mar 2020).

21. Geijer H, Udumyan R, Lohse G, Nilsagård Y. Temperature measurements with a temporal scanner: systematic review and meta-analysis. *BMJ Open* 2016; **6(3):** e009509.

22. Niven DJ, Gaudet JE, Laupland KB, *et al.* Accuracy of peripheral thermometers for estimating temperature: a systematic review and meta-analysis. *Ann Intern Med* 2015; **163(10):** 768–777.

23. Craig JV, Lancaster GA, Williamson PR, Smyth RL. Temperature measured at the axilla compared with rectum in children and young people: systematic review. *BMJ* 2000; **320(7243):** 1174–1178.

24. Zhen C, Xia Z, Long L, Pu Y. Accuracy of infrared ear thermometry in children: a meta-analysis and systematic review. *Clin Pediatr (Phila)* 2014; **53(12):** 1158–1165.

25. Zhen C, Xia Z, Ya Jun Z, *et al.* Accuracy of infrared tympanic thermometry used in the diagnosis of fever in children: a systematic review and meta-analysis. *Clin Pediatr (Phila)* 2015; **54(2):** 114–126.

26. Apa H, Gözmen S, Keskin-Gözmen Ş, *et al.* Clinical accuracy of non-contact infrared thermometer from umbilical region in children: a new side. *Turk J Pediatr* 2016; **58(2):** 180–186.

27. Kelly M, Sahm LJ, Shiely F, *et al.* Parental knowledge, attitudes and beliefs on fever: a cross-sectional study in Ireland. *BMJ Open* 2017; **7(7):** e015684.

28. Kelly M, McCarthy S, O'Sullivan R, *et al.* Drivers for inappropriate fever management in children: a systematic review. *Int J Clin Pharm* 2016; **38(4):** 761–770.

29. Kai J. Parents' perceptions of taking babies' rectal temperature. *BMJ* 1993; **307(6905):** 660–662.