

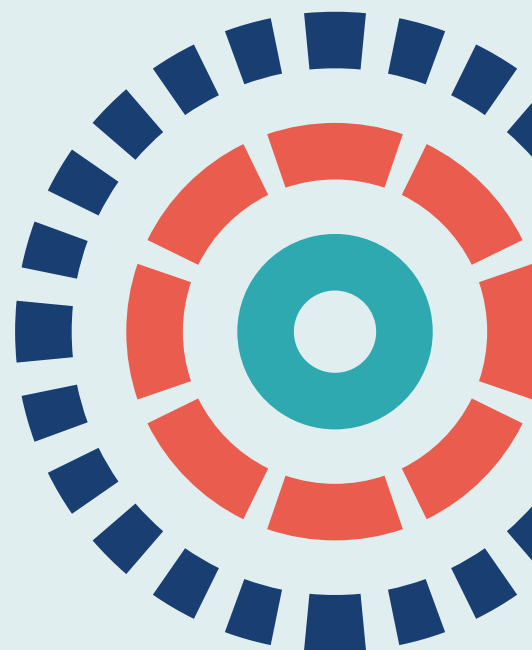
## Health Technology Assessment

Volume 24 • Issue 53 • October 2020

ISSN 1366-5278

# Non-contact infrared thermometers compared with current approaches in primary care for children aged 5 years and under: a method comparison study

*Ann Van den Bruel, Jan Verbakel, Kay Wang, Susannah Fleming, Gea Holtman, Margaret Glogowska, Elizabeth Morris, George Edwards, Fatene Abakar Ismail, Kathryn Curtis, James Goetz, Grace Barnes, Ralitsa Slivkova, Charlotte Nesbitt, Suhail Aslam, Ealish Swift, Harriet Williams and Gail Hayward*





# Non-contact infrared thermometers compared with current approaches in primary care for children aged 5 years and under: a method comparison study

Ann Van den Bruel,<sup>1,2\*</sup> Jan Verbakel,<sup>1,2</sup> Kay Wang<sup>id</sup>,<sup>1</sup>  
Susannah Fleming<sup>id</sup>,<sup>1</sup> Gea Holtman<sup>id</sup>,<sup>1,3</sup>  
Margaret Glogowska<sup>id</sup>,<sup>1</sup> Elizabeth Morris<sup>id</sup>,<sup>1</sup>  
George Edwards<sup>id</sup>,<sup>1</sup> Fatene Abakar Ismail<sup>id</sup>,<sup>1</sup>  
Kathryn Curtis<sup>id</sup>,<sup>1</sup> James Goetz<sup>id</sup>,<sup>1</sup> Grace Barnes<sup>id</sup>,<sup>1</sup>  
Ralitsa Slivkova<sup>id</sup>,<sup>1</sup> Charlotte Nesbitt<sup>id</sup>,<sup>1</sup> Suhail Aslam<sup>id</sup>,<sup>1</sup>  
Ealish Swift<sup>id</sup>,<sup>1</sup> Harriet Williams<sup>id</sup><sup>1</sup> and Gail Hayward<sup>id</sup><sup>1</sup>

<sup>1</sup>Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK

<sup>2</sup>Academic Centre for Primary Care, University of Leuven, Leuven, Belgium

<sup>3</sup>Department of General Practice and Elderly Care Medicine, University Medical Centre Groningen, University of Groningen, Groningen, the Netherlands

\*Corresponding author

**Declared competing interests of authors:** Ann Van den Bruel was a member of the Health Technology Assessment (HTA) Maternal, Neonatal and Child Health panel, and was a member of the Diagnosis and Screening Methods group from 2015 to 2018. Gail Hayward was a member of the HTA Commissioning Board. Susannah Fleming was funded under a Programme Grants for Applied Research programme grant with number RP-PG-1210-12003 [Monitoring Long-term Conditions in Primary Care; URL: [www.journals.library.nihr.ac.uk/programmes/pgfar/rp-pg-1210-12003](http://www.journals.library.nihr.ac.uk/programmes/pgfar/rp-pg-1210-12003) (accessed June 2020)] while working on this report.

Published October 2020

DOI: 10.3310/hta24530

This report should be referenced as follows:

Van den Bruel A, Verbakel J, Wang K, Fleming S, Holtman G, Glogowska M, *et al.* Non-contact infrared thermometers compared with current approaches in primary care for children aged 5 years and under: a method comparison study. *Health Technol Assess* 2020;**24**(53).

*Health Technology Assessment* is indexed and abstracted in *Index Medicus/MEDLINE*, *Excerpta Medica/EMBASE*, *Science Citation Index Expanded (SciSearch®)* and *Current Contents®/Clinical Medicine*.



ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 3.370

*Health Technology Assessment* is indexed in MEDLINE, CINAHL, EMBASE, the Cochrane Library and Clarivate Analytics Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) ([www.publicationethics.org/](http://www.publicationethics.org/)).

Editorial contact: [journals.library@nihr.ac.uk](mailto:journals.library@nihr.ac.uk)

The full HTA archive is freely available to view online at [www.journalslibrary.nihr.ac.uk/hta](http://www.journalslibrary.nihr.ac.uk/hta). Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: [www.journalslibrary.nihr.ac.uk](http://www.journalslibrary.nihr.ac.uk)

## Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

## HTA programme

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

## This report

The research reported in this issue of the journal was funded by the HTA programme as project number 16/45/01. The contractual start date was in December 2016. The draft report began editorial review in April 2019 and was accepted for publication in October 2019. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care.

© Queen's Printer and Controller of HMSO 2020. This work was produced by Van den Bruel *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by the NIHR Journals Library ([www.journalslibrary.nihr.ac.uk](http://www.journalslibrary.nihr.ac.uk)), produced by Prepress Projects Ltd, Perth, Scotland ([www.prepress-projects.co.uk](http://www.prepress-projects.co.uk)).

## Editor-in-Chief of *Health Technology Assessment* and NIHR Journals Library

---

**Professor Ken Stein** Professor of Public Health, University of Exeter Medical School, UK

## NIHR Journals Library Editors

---

**Professor John Powell** Chair of HTA and EME Editorial Board and Editor-in-Chief of HTA and EME journals. Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK, and Senior Clinical Researcher, Nuffield Department of Primary Care Health Sciences, University of Oxford, UK

**Professor Andrée Le May** Chair of NIHR Journals Library Editorial Group (HS&DR, PGfAR, PHR journals) and Editor-in-Chief of HS&DR, PGfAR, PHR journals

**Professor Matthias Beck** Professor of Management, Cork University Business School, Department of Management and Marketing, University College Cork, Ireland

**Dr Tessa Crilly** Director, Crystal Blue Consulting Ltd, UK

**Dr Eugenia Cronin** Senior Scientific Advisor, Wessex Institute, UK

**Dr Peter Davidson** Consultant Advisor, Wessex Institute, University of Southampton, UK

**Ms Tara Lamont** Senior Scientific Adviser (Evidence Use), Wessex Institute, University of Southampton, UK

**Dr Catriona McDaid** Senior Research Fellow, York Trials Unit, Department of Health Sciences, University of York, UK

**Professor William McGuire** Professor of Child Health, Hull York Medical School, University of York, UK

**Professor Geoffrey Meads** Professor of Wellbeing Research, University of Winchester, UK

**Professor John Norrie** Chair in Medical Statistics, University of Edinburgh, UK

**Professor James Raftery** Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

**Dr Rob Riemsma** Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

**Professor Helen Roberts** Professor of Child Health Research, UCL Great Ormond Street Institute of Child Health, UK

**Professor Jonathan Ross** Professor of Sexual Health and HIV, University Hospital Birmingham, UK

**Professor Helen Snooks** Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

**Professor Ken Stein** Professor of Public Health, University of Exeter Medical School, UK

**Professor Jim Thornton** Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

**Professor Martin Underwood** Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, UK

Please visit the website for a list of editors: [www.journalslibrary.nihr.ac.uk/about/editors](http://www.journalslibrary.nihr.ac.uk/about/editors)

**Editorial contact:** [journals.library@nihr.ac.uk](mailto:journals.library@nihr.ac.uk)

# Abstract

## Non-contact infrared thermometers compared with current approaches in primary care for children aged 5 years and under: a method comparison study

Ann Van den Bruel,<sup>1,2\*</sup> Jan Verbakel,<sup>1,2</sup> Kay Wang<sup>1</sup>,  
Susannah Fleming<sup>1</sup>, Gea Holtman<sup>1,3</sup>, Margaret Glogowska<sup>1</sup>,  
Elizabeth Morris<sup>1</sup>, George Edwards<sup>1</sup>, Fatene Abakar Ismail<sup>1</sup>,  
Kathryn Curtis<sup>1</sup>, James Goetz<sup>1</sup>, Grace Barnes<sup>1</sup>, Ralitsa Slivkova<sup>1</sup>,  
Charlotte Nesbitt<sup>1</sup>, Suhail Aslam<sup>1</sup>, Ealish Swift<sup>1</sup>, Harriet Williams<sup>1</sup>  
and Gail Hayward<sup>1</sup>

<sup>1</sup>Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK

<sup>2</sup>Academic Centre for Primary Care, University of Leuven, Leuven, Belgium

<sup>3</sup>Department of General Practice and Elderly Care Medicine, University Medical Centre Groningen, University of Groningen, Groningen, the Netherlands

\*Corresponding author [ann.vandenbruel@kuleuven.be](mailto:ann.vandenbruel@kuleuven.be)

**Background:** Current options for temperature measurement in children presenting to primary care include either electronic axillary or infrared tympanic thermometers. Non-contact infrared thermometers could reduce both the distress of the child and the risk of cross-infection.

**Objectives:** The objective of this study was to compare the use of non-contact thermometers with the use of electronic axillary and infrared tympanic thermometers in children presenting to primary care.

**Design:** Method comparison study with a nested qualitative study.

**Setting:** Primary care in Oxfordshire.

**Participants:** Children aged  $\leq 5$  years attending with an acute illness.

**Interventions:** Two types of non-contact infrared thermometers [i.e. Thermofocus (Tecnimed, Varese, Italy) and Firhealth (Firhealth, Shenzhen, China)] were compared with an electronic axillary thermometer and an infrared tympanic thermometer.

**Main outcome measures:** The primary outcome was agreement between the Thermofocus non-contact infrared thermometer and the axillary thermometer. Secondary outcomes included agreement between all other sets of thermometers, diagnostic accuracy for detecting fever, parental and child ratings of acceptability and discomfort, and themes arising from our qualitative interviews with parents.

**Results:** A total of 401 children (203 boys) were recruited, with a median age of 1.6 years (interquartile range 0.79–3.38 years). The readings of the Thermofocus non-contact infrared thermometer differed from those of the axillary thermometer by  $-0.14^{\circ}\text{C}$  (95% confidence interval  $-0.21$  to  $-0.06^{\circ}\text{C}$ ) on average with the lower limit of agreement being  $-1.57^{\circ}\text{C}$  (95% confidence interval  $-1.69$  to  $-1.44^{\circ}\text{C}$ ) and the upper limit being  $1.29^{\circ}\text{C}$  (95% confidence interval  $1.16$  to  $1.42^{\circ}\text{C}$ ). The readings of the Firhealth non-contact infrared thermometer differed from those of the axillary thermometer by  $-0.16^{\circ}\text{C}$  (95% confidence interval  $-0.23$  to  $-0.09^{\circ}\text{C}$ ) on average, with the lower limit of agreement

being  $-1.54^{\circ}\text{C}$  (95% confidence interval  $-1.66$  to  $-1.41^{\circ}\text{C}$ ) and the upper limit being  $1.22^{\circ}\text{C}$  (95% confidence interval  $1.10$  to  $1.34^{\circ}\text{C}$ ). The difference between the first and second readings of the Thermofocus was  $-0.04^{\circ}\text{C}$  (95% confidence interval  $-0.07$  to  $-0.01^{\circ}\text{C}$ ); the lower limit was  $-0.56^{\circ}\text{C}$  (95% confidence interval  $-0.60$  to  $-0.51^{\circ}\text{C}$ ) and the upper limit was  $0.47^{\circ}\text{C}$  (95% confidence interval  $0.43$  to  $0.52^{\circ}\text{C}$ ). The difference between the first and second readings of the Firhealth thermometer was  $0.01^{\circ}\text{C}$  (95% confidence interval  $-0.02$  to  $0.04^{\circ}\text{C}$ ); the lower limit was  $-0.60^{\circ}\text{C}$  (95% confidence interval  $-0.65$  to  $-0.54^{\circ}\text{C}$ ) and the upper limit was  $0.61^{\circ}\text{C}$  (95% confidence interval  $0.56$  to  $0.67^{\circ}\text{C}$ ). Sensitivity and specificity for the Thermofocus non-contact infrared thermometer were 66.7% (95% confidence interval 38.4% to 88.2%) and 98.0% (95% confidence interval 96.0% to 99.2%), respectively. For the Firhealth non-contact infrared thermometer, sensitivity was 12.5% (95% confidence interval 1.6% to 38.3%) and specificity was 99.4% (95% confidence interval 98.0% to 99.9%). The majority of parents found all methods to be acceptable, although discomfort ratings were highest for the axillary thermometer. The non-contact thermometers required fewer readings than the comparator thermometers.

**Limitations:** A method comparison study does not compare new methods against a reference standard, which in this case would be central thermometry requiring the placement of a central line, which is not feasible or acceptable in primary care. Electronic axillary and infrared tympanic thermometers have been found to have moderate agreement themselves with central temperature measurements.

**Conclusions:** The 95% limits of agreement are  $> 1^{\circ}\text{C}$  for both non-contact infrared thermometers compared with electronic axillary and infrared tympanic thermometers, which could affect clinical decision-making. Sensitivity for fever was low to moderate for both non-contact thermometers.

**Future work:** Better methods for peripheral temperature measurement that agree well with central thermometry are needed.

**Trial registration:** Current Controlled Trials ISRCTN15413321.

**Funding:** This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 24, No. 53. See the NIHR Journals Library website for further project information.



# Contents

List of tables	ix
List of figures	xi
List of abbreviations	xiii
Plain English summary	xv
Scientific summary	xvii
<b>Chapter 1</b> Introduction	<b>1</b>
<b>Chapter 2</b> Methods	<b>3</b>
Sample size	4
Protocol changes	4
Analyses	4
Qualitative data collection and analysis	5
<i>Recruitment</i>	5
<i>Data collection</i>	5
<i>Data analysis</i>	5
<b>Chapter 3</b> Results	<b>7</b>
Acceptability of the thermometers	9
Qualitative interviews	10
<i>Electronic axillary thermometer</i>	11
<i>Infrared tympanic thermometer</i>	12
<i>Non-contact infrared thermometers</i>	12
<b>Chapter 4</b> Patient and public involvement	<b>15</b>
<b>Chapter 5</b> Discussion	<b>17</b>
Summary of the main findings	17
Strengths and limitations	17
Comparison with existing literature	18
Implications for research and practice	18
<b>Acknowledgements</b>	<b>21</b>
<b>References</b>	<b>23</b>
<b>Appendix 1</b> Participant characteristics	<b>25</b>
<b>Appendix 2</b> Bland–Altman plots	<b>27</b>



# List of tables

<b>TABLE 1</b> Participant characteristics	<b>7</b>
<b>TABLE 2</b> Method agreement results	<b>9</b>
<b>TABLE 3</b> Diagnostic accuracy for fever defined as $\geq 38^{\circ}\text{C}$ by the electronic axillary thermometer	<b>9</b>
<b>TABLE 4</b> Number of attempts for each thermometer	<b>10</b>
<b>TABLE 5</b> Summary of main themes of the qualitative interviews	<b>11</b>
<b>TABLE 6</b> Qualitative interview participant characteristics	<b>25</b>



## List of figures

<b>FIGURE 1</b> Bland–Altman plot for agreement between the Thermofocus non-contact thermometer and the electronic axillary thermometer	<b>8</b>
<b>FIGURE 2</b> Bland–Altman plot for agreement between the Firhealth non-contact thermometer and the electronic axillary thermometer	<b>8</b>
<b>FIGURE 3</b> Results of the Patient Discomfort Scale	<b>10</b>
<b>FIGURE 4</b> Bland–Altman plots of agreement for the other comparisons	<b>27</b>



## List of abbreviations

CE	Conformité Européenne	NCIT	non-contact infrared thermometer
CI	confidence interval	NICE	National Institute for Health and Care Excellence
FDA	Food and Drug Administration	NIHR	National Institute for Health Research
GP	general practitioner	PPI	patient and public involvement
IQR	interquartile range		





## Plain English summary

General practitioners commonly measure children's temperature using a thermometer placed in the armpit or ear canal. New 'non-contact' thermometers use infrared light to measure temperature without touching the child. They are easy to use and there is no risk of passing on infections. However, we do not know how well they measure temperature compared with thermometers that use the armpit or the ear.

This study aimed to compare two non-contact thermometers with current thermometers. We measured children's temperature with all thermometer types, and asked children and their parents about their views.

The study was performed in general practices in Oxfordshire with children aged  $\leq 5$  years who had come to see their general practitioner because they had recently become unwell.

Both the cheaper and more expensive non-contact thermometers gave slightly lower temperature readings on average than current thermometers. The vast majority of readings ranged from 1.6 °C lower to 1.3 °C higher than current thermometers. The detection of fever of at least 38 °C was low to moderate for both non-contact thermometers.

Most parents did not think that their child was distressed by having their temperature taken using any of the thermometers, but the armpit thermometer was rated as the least comfortable. When interviewed, parents were more negative about the armpit thermometers, although still willing to use them if they were recommended by doctors.

Although we found that the readings from the different thermometers did not match, we do not know whether the non-contact or the current thermometers were giving readings that were closer to the real temperature of the child. To understand this, we would need to do a study that included a more invasive procedure for temperature assessment.



# Scientific summary

## Background

Acute infections are very common in children, resulting in a high proportion of children seeking health care for acute infections. There is specific guidance for management of febrile illness in children aged  $\leq 5$  years, including the measurement of temperature in each child presenting with fever symptoms using either electronic axillary thermometers or infrared tympanic thermometers in children aged  $\geq 4$  weeks. However, axillary thermometers require the child to be at least partially undressed and the thermometer has to be held in place for at least 30 seconds. Infrared tympanic thermometers are easier to use but may be inaccurate when the ear drum is insufficiently exposed because of ear wax or insufficient straightening of the ear canal. Non-contact infrared thermometers could reduce both the distress of the child and the risk of cross-infection. These thermometers do not require direct contact with the child, are fast and do not require additional disposable probe covers. This means that they could result in less discomfort, while minimising the risk of cross-infection. However, our Horizon Scan of this diagnostic technology found limited evidence of comparative accuracy to standard methods.

## Objectives

The objectives of this study were to compare non-contact thermometers with electronic axillary and infrared tympanic thermometers in children presenting to primary care. These objectives included an assessment of agreement between different thermometer types, and exploration of the acceptability and discomfort to patients.

## Design

This was a method comparison study with a nested qualitative study.

## Setting

The study was conducted in nine general practices and one out-of-hours centre in Oxfordshire.

## Participants

Children aged  $\leq 5$  years attending with an acute illness of a maximum of 14 days were eligible. Children for whom acute trauma was the main reason for presentation, who were clinically unstable, who had already been enrolled in the study or whose parents were unable to understand trial material in English were excluded from the study.

## Interventions

Each child had their temperature measured using four different thermometers:

1. electronic axillary – Welch Allyn SureTemp (Welch Allyn®, USA)
2. infrared tympanic – Braun Thermoscan (Braun GmbH, Kronberg, Germany) (not in children  $< 4$  weeks of age)
3. non-contact infrared thermometer 1 – Thermofocus 0800 (Tecnimed Srl, Varese, Italy)
4. non-contact infrared thermometer 2 – Firhealth Forehead Thermometer (Firhealth, Shenzhen, China).

The Thermofocus non-contact infrared thermometer was included as it was the most extensively evaluated thermometer in other settings. The Firhealth device was included as an example of a cheaper, non-contact infrared thermometer. The order in which the thermometers were used was randomised prior to the study start for each participant [using a random number generator ([www.random.org](http://www.random.org))]. Measurements were performed consecutively; a second measurement with each non-contact thermometer was carried out to evaluate reproducibility. Failed measurements were recorded.

We assessed discomfort to the child using the Patient Discomfort Scale and Wong-Baker FACES® Pain Rating Scale (Wong-Baker FACES® Foundation, Oklahoma City, OK, USA). Parents were also asked to score acceptability of each thermometer on a visual analogue scale and rank thermometers by preference.

## Analysis

The sample size calculation was based on a desired accuracy of  $\pm 0.075$  °C for the 95% confidence intervals of the limits of agreement, and the standard deviation of the agreement between temperatures measured by non-contact and electronic axillary thermometers would be 0.5 °C. The sample size initially set at 533 participants was revised during the course of the study to 400 participants, assuming that 0.10 °C accuracy would be sufficient. Analyses of agreement were conducted based on Bland-Altman plots, which provide an indication of bias and limits of agreement between the measurements. Exact 95% confidence intervals around this estimate have been calculated. Diagnostic accuracy for detecting fever (temperature of  $\geq 38$  °C measured by the electronic axillary thermometer) was analysed by calculating sensitivity, specificity, predictive values and likelihood ratios, with 95% confidence intervals. Other analyses, such as failure rates and indeterminate readings, are reported as proportions with their corresponding estimates of precision. The scores on the visual analogue scale and the Patient Discomfort Scale have been analysed using non-parametric techniques, resulting in median acceptability (and interquartile ranges) for each thermometer.

For the nested qualitative study, parents were purposively sampled to achieve maximum variation in gender, age of parent, age of child, ethnicity and number of other children in the household. Recruitment continued until the research team agreed that data saturation had been achieved and sufficient explanation for the categories generated was reached. Interviews were semistructured and by telephone ( $n = 20$ ) or face to face ( $n = 1$ ) following a flexible topic guide developed by the research team and patient and public involvement panel, which evolved in response to emerging themes. Data analysis followed a thematic approach with the assistance of NVivo (version 11; QSR International, Warrington, UK). 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 (EM and MG). The codes and themes were developed and interpreted in discussion with the wider research team.

## Main outcome measures

The primary outcome was agreement between the Thermofocus thermometer and the axillary thermometer. Secondary outcomes included agreement between all other sets of thermometers, diagnostic accuracy for detecting fever, and acceptability and discomfort.

## Results

A total of 401 children (203 boys) were recruited, with a median age of 1.6 years (interquartile range 0.79–3.38 years). Most children were of white British ethnicity (69.83%). Approximately 30% of the children were feverish at the time of inclusion. There were 396 temperature readings with the Thermofocus non-contact thermometer (first measurement), 399 with the Firhealth non-contact

thermometer (first measurement), 376 with the electronic axillary thermometer and 390 with the tympanic thermometer. Second measurements with the Thermofocus and Firhealth non-contact thermometers resulted in 395 and 397 readings, respectively.

On average, the non-contact thermometers showed lower readings than the electronic axillary thermometer. The readings of the Thermofocus differed from those of the electronic axillary thermometer by a mean of  $-0.14^{\circ}\text{C}$  (95% confidence interval  $-0.21$  to  $-0.06^{\circ}\text{C}$ ), with the lower limit of agreement being  $-1.57^{\circ}\text{C}$  (95% confidence interval  $-1.69$  to  $-1.44^{\circ}\text{C}$ ) and the upper limit being  $1.29^{\circ}\text{C}$  (95% confidence interval  $1.16$  to  $1.42^{\circ}\text{C}$ ). The readings of the Firhealth non-contact infrared thermometer differed from those of the electronic axillary thermometer by a mean of  $-0.16^{\circ}\text{C}$  (95% confidence interval  $-0.23$  to  $-0.09^{\circ}\text{C}$ ), with the lower limit of agreement being  $-1.54^{\circ}\text{C}$  (95% confidence interval  $-1.66$  to  $-1.41^{\circ}\text{C}$ ) and the upper limit being  $1.22^{\circ}\text{C}$  (95% confidence interval  $1.10$  to  $1.34^{\circ}\text{C}$ ).

Agreement between the first and second readings with each non-contact thermometer resulted in a mean difference of  $-0.04^{\circ}\text{C}$  (95% confidence interval  $-0.07$  to  $-0.01^{\circ}\text{C}$ ), a lower limit of agreement of  $-0.56^{\circ}\text{C}$  (95% confidence interval  $-0.60$  to  $-0.51^{\circ}\text{C}$ ) and an upper limit of agreement of  $0.47^{\circ}\text{C}$  (95% confidence interval  $0.43$  to  $0.52^{\circ}\text{C}$ ) for the Thermofocus. There was a mean difference of  $0.01^{\circ}\text{C}$  (95% confidence interval  $-0.02$  to  $0.04^{\circ}\text{C}$ ), a lower limit of agreement of  $-0.60^{\circ}\text{C}$  (95% confidence interval  $-0.65$  to  $-0.54^{\circ}\text{C}$ ) and an upper limit of agreement of  $0.61^{\circ}\text{C}$  (95% confidence interval  $0.56$  to  $0.67^{\circ}\text{C}$ ) for the Firhealth thermometer.

Sensitivity and specificity of the Thermofocus non-contact thermometer were 66.7% (95% confidence interval 38.4% to 88.2%) and 98.0% (95% confidence interval 96.0% to 99.2%), respectively. For the Firhealth thermometer, sensitivity was 12.5% (95% confidence interval 1.6% to 38.3%) and specificity was 99.4% (95% confidence interval 98.0% to 99.9%). Similarly, the sensitivity of the tympanic thermometer to detect fever defined by axillary temperature measurement was 62.5% (95% confidence interval 35.4% to 84.8%) and specificity was 96.0% (95% confidence interval 93.4% to 97.8%).

The non-contact thermometers required fewer attempts to obtain a reading than the axillary and tympanic thermometers. In addition, there were nine technical failures with the Thermofocus non-contact thermometer, one with the Firhealth non-contact thermometer, eight with the electronic axillary thermometer and 10 with the tympanic thermometer.

The majority of parents found all methods acceptable, although discomfort ratings were highest for the axillary thermometer. Median parental acceptability as assessed with a visual analogue scale was highest for the Firhealth non-contact thermometer (8.23 cm, 95% confidence interval 8.04 to 8.41 cm), followed by the Thermofocus non-contact thermometer (7.82 cm, 95% confidence interval 7.60 to 8.03 cm), the tympanic thermometer (7.12 cm, 95% confidence interval 6.89 to 7.34 cm) and, finally, the electronic axillary thermometer (5.01 cm, 95% confidence interval 4.71 to 5.31 cm).

We interviewed 21 parents. Parents' experiences with the axillary thermometers were mostly described in negative language, such as being uncomfortable and impractical, whereas experiences with the tympanic thermometer were more neutral. Parents were pleasantly surprised by the practicality and convenience of the non-contact thermometers, which they had been unfamiliar with until then.

## Limitations

The design of a method comparison study does not compare new methods against a gold standard, which in this case would be central thermometry requiring the placement of a central line, which is not feasible in primary care. Digital and tympanic thermometers have been found to have moderate agreement themselves with central temperature measurements.

## Conclusions

The two types of non-contact infrared thermometers evaluated in this study on average resulted in lower temperature readings than the currently recommended approaches to temperature measurement in children presenting to the general practitioner with acute illness (i.e. electronic axillary and infrared tympanic thermometers). The mean difference was  $-0.10$  to  $-0.16$  °C, and the lower and upper limits of agreement ranged from  $-1.47$  to  $-1.57$  °C and  $1.22$  to  $1.35$  °C, respectively. Sensitivity for fever was low to moderate in both cases.

## Future work

Better methods for peripheral temperature measurement that agree well with central thermometry are needed.

## Trial registration

This trial is registered as ISRCTN15413321.

## Funding

This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 24, No. 53. See the NIHR Journals Library website for further project information.

# Chapter 1 Introduction

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

Guidelines recommend the measurement of temperature in each child presenting with fever symptoms using either electronic axillary thermometers or infrared tympanic thermometers in children aged  $\geq 4$  weeks.<sup>4</sup> However, axillary thermometers require health-care professionals to undress the child and hold the thermometer in the axilla for at least 30 seconds.<sup>5</sup> Infrared tympanic thermometers are easier to use, but may be inaccurate owing to ear wax or insufficient straightening of the ear canal.<sup>6</sup> 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 or distress, measurement without interrupting sleep, minimal risk of cross-infection and no requirement for additional disposable probe covers.<sup>7</sup>

This study was set up in response to a primary care diagnostic technology update by our group, as part of our Horizon Scan programme within the National Institute for Health Research (NIHR) Diagnostic Evidence Cooperative Oxford. The results of the review were published in November 2013 as an open-access publication on our website<sup>8</sup> and published in the *British Journal of General Practice* by Wang *et al.*<sup>7</sup> All Horizon Scans are conducted according to standardised methods. The search strategy includes systematic searches of MEDLINE, EMBASE, MEDION, The Cochrane Library, TRIP and NHS Evidence, using tailored search strategies for each topic and limiting to English-language publications. For guidelines, we searched the National Institute for Health and Care Excellence (NICE)'s, the Scottish Intercollegiate Guidelines Network (SIGN)'s and the relevant professional bodies' websites. We searched for supplementary information on manufacturer/trade websites and through web search engines. When assessing publications for inclusion in reports, we used critical appraisal tools developed by the Centre for Evidence Based Medicine, which are available at [www.cebm.net/index.aspx?o=1157](http://www.cebm.net/index.aspx?o=1157) (accessed December 2019). Reports are reviewed and edited by at least two general practitioners (GPs) and a pathology commissioner, and health economics sections are written by a health economist.

The report on non-contact thermometers summarised evidence on seven different non-contact thermometers in different price ranges. The Thermofocus (Tecnimed, Varese, Italy) was the most extensively studied thermometer (five out of six studies) for agreement with mercury-in-glass,<sup>9,10</sup> electronic rectal,<sup>11,12</sup> electronic axillary,<sup>12,13</sup> electronic oral,<sup>12</sup> infrared tympanic<sup>13</sup> and temporal artery thermometers.<sup>13</sup> Some studies had more than one comparison. Agreement was found to be moderate to low. In addition, four studies were identified that assessed non-contact thermometer accuracy in detecting fever. Fever was defined as a temperature reading of  $\geq 38$  °C, measured by a mercury-in-glass thermometer axillary<sup>9</sup> or rectally,<sup>10</sup> a tympanic thermometer<sup>14</sup> or an electronic thermometer used in the axilla, orally or rectally.<sup>12</sup> These studies found that non-contact thermometers had high sensitivity and specificity for detecting fever, with sensitivities ranging from 77% to 97% and specificities ranging from 75% to 97%. There was also evidence based on one study that children found the non-contact thermometers more acceptable than a mercury-in-glass axillary thermometer.

To our knowledge, there have been no other systematic reviews on non-contact thermometers apart from our Horizon Scan review.<sup>7</sup> However, there are several systematic reviews on other thermometer types, all highlighting the poor performance of electronic axillary and infrared tympanic thermometers compared with core body temperature measurements.<sup>6,16,17</sup> Compared with central thermometry, which ideally is measured via a pulmonary artery catheter, measurements using a tympanic membrane,

temporal artery and axillary thermometer were all found to have limits of agreement wider than 0.5 °C in children in a 2015 systematic review including 75 studies.<sup>16</sup> In contrast, rectal measurements were found to have much lower limits of agreement of -0.24 to 0.04 °C.

Specifically comparing rectal with axillary readings using the same device, Craig *et al.*<sup>17</sup> found in their systematic review of 37 papers with > 5000 children that the mean axillary temperature was always lower than the mean rectal temperature, and the difference was larger for electronic thermometers than for mercury-in-glass thermometers [i.e. the pooled mean difference of rectal minus axillary temperature was 0.25 °C (95% limits of agreement -0.15 to 0.65 °C) for mercury thermometers and 0.85 °C (95% limits of agreement -0.19 to 1.90 °C) for electronic thermometers].<sup>17</sup> In another systematic review<sup>6</sup> including 44 papers with almost 6000 children, the same authors compared tympanic thermometers with rectal temperature using mercury, electronic or indwelling probe thermometers. The pooled mean difference between rectal and tympanic readings was 0.29 °C (95% limits of agreement -0.74 to 1.32 °C).

Reports of agreement between NCITs and conventional thermometers have been variable. Only one study has reported comparable data for the Thermofocus brand compared with an axillary measurement using a mercury-in-glass thermometer department<sup>9</sup> recruiting 251 children aged between 1 month and 18 months in emergency clinics, paediatric outpatient, and one primary care centre. They found an overall mean difference of 0.07 °C and limits of agreement of -0.62 °C (95% CI -0.67 to -0.47 °C) and 0.76 °C (95% CI 0.61 to 0.91 °C).

However, comparisons with infrared tympanic thermometers reported larger mean differences of -0.38 °C [95% confidence interval (CI) -1.47 to 0.70 °C]<sup>18</sup> and 2.34 °C (95% CI 0.26 to 4.42 °C).<sup>14</sup> Comparisons with rectal thermometry have also yielded variable results, with mean differences reported of 0.029<sup>10</sup> and 0.34 °C.<sup>19</sup> Different devices have been shown to perform differently: closer agreement is reported with the Thermofocus than with the Beurer.<sup>20</sup>

Finally, although NCITs are mostly reported to have high sensitivity and specificity in detecting a fever of  $\geq 38$  °C using conventional thermometry,<sup>9,10,12,14</sup> sensitivity was estimated as only 27%<sup>19</sup> and 12%<sup>20</sup> in two studies.

In conclusion, agreement between non-contact thermometers and other methods of thermometry has been found to be variable and highly dependent on the thermometer with which the NCITs are compared. In addition to the lack of clear conclusions from existing studies, there is a lack of generalisability of these data to primary care settings as most previous studies were conducted in paediatric inpatient populations<sup>14,19,20</sup> or mixed hospital ambulatory care and ward settings.<sup>10,13</sup> Furthermore, NCITs have been mainly compared with thermometers that are not currently recommended for use in children, including rectal<sup>10,11,19,20</sup> and mercury-in-glass axillary thermometers.<sup>9</sup>

Understanding the performance of NCITs in a primary care paediatric population compared with infrared tympanic and electronic axillary thermometers could allow introduction of this new technology into routine practice. We evaluated the agreement between two NCIT models at different price points versus electronic axillary and infrared tympanic thermometers in children who present with acute illness in primary care.



## Chapter 2 Methods

This was a cross-sectional method agreement study with a nested qualitative study. The study ran from April 2017 to August 2018.

Children aged up to 5 years with an acute illness of a maximum of 14 days presenting to a general practice or an out-of-hours service were eligible for inclusion. Children for whom acute trauma was the main reason for presentation, who were clinically unstable, who had already been enrolled in the study or whose parents were unable to understand trial material in English were excluded from the study.

The study ran in nine general practices and one out-of-hours centre in Oxfordshire (UK). All potentially eligible children and their parents were consecutively approached by a research assistant from the study team in the waiting room for possible recruitment to the study. There were 11 research assistants involved in the study, all of whom had been trained in study procedures by the research team, which included an off-site explanation and dry run of study procedures, and on-site shadowing. Parents were handed patient information sheets to explain the goal of the study, the risks and the procedures. After verbal informed consent was obtained, temperature measurements were conducted either prior to or after the child's appointment with the GP. The research assistant recorded demographic information (including household composition, parental age and ethnicity) and baseline information (including prior fever medication use, parental impression of fever and fever duration) for each child.

The child had their temperature measured using four different thermometers:

1. electronic axillary – Welch Allyn SureTemp (Welch Allyn®, USA)
2. infrared tympanic – Braun Thermoscan (Braun GmbH, Kronberg, Germany)
3. NCIT 1 – Thermofocus 0800 (Tecnimed Srl, Varese, Italy)
4. NCIT 2 – Firhealth Forehead Thermometer (Firhealth, Shenzhen, China).

The Thermofocus NCIT was included as it was the most extensively evaluated thermometer in other settings. The Firhealth device was included as an example of a cheaper NCIT that was easily available for purchase. At the time of writing this report, the Thermofocus was priced at around £80.00 and the Firhealth at around £30.00. The Thermofocus 0800H5 by Thermofocus is a certified medical device, Food and Drug Administration (FDA) approved and Conformité Européenne (CE) marked. It is a NCIT that measures body temperature by pointing the thermometer at the centre of the forehead. The optimal distance between the forehead and the thermometer is determined by a light-emitting diode (LED) system that emits two light beams. As the thermometer is moved closer to the forehead, at the right distance (approximately 3 cm) the two beams converge to a single red dot. Releasing the button until the lights flash will result in a temperature reading. A manual calibration system is provided with the thermometer to improve accurate readings. No extensive training is needed to operate the device and a comprehensive user manual is provided with the device.

The Firhealth Forehead Thermometer by Firhealth is a FDA-approved, CE-marked NCIT that works in essentially the same way as the Thermofocus. To measure temperature, the thermometer probe is pointed at the forehead at a distance of 1–5 cm. After pressing the 'Measure' button, the device beeps and displays the temperature reading on the liquid-crystal display (LCD) screen.

The comparator thermometers (i.e. the electronic axillary and infrared tympanic thermometer) were chosen because these are recommended by NICE for use in acutely ill children. Children < 4 weeks of age did not have the tympanic thermometer measurement as per NICE recommendation; as a result, we opted for the electronic axillary thermometer for our primary outcome. The order in which the thermometers were used was randomised by our statistician prior to the study start for each participant (using a random number generator: [www.random.org](http://www.random.org)).

Measurements were performed consecutively in the shortest time frame possible, and no medication or drinks were administered between measurements. Once the four primary measurements were complete, a second measurement was taken with each NCIT to evaluate reproducibility. We also recorded failed measurements due to lack of co-operation of the child after three attempts, mechanical issues (operational or technological failure) and clinically implausible readings (based on researcher's assessment).

Children's reaction to the different measurements was recorded by the Patient Discomfort Scale<sup>21</sup> as assessed by their parent(s). Children aged 4 or 5 years also completed the Wong-Baker FACES® Pain Rating Scale (Wong-Baker FACES® Foundation, Oklahoma City, OK, USA), which scores pain from 0 to 10 in 2-point increments.<sup>22</sup> Parents were asked to score the acceptability of each thermometer on a visual analogue scale by indicating their acceptability of each thermometer on a scale of 0–10 cm, and rank the thermometers by preference. To avoid bias, parents and children were blinded to the temperature measurements until they rated their acceptability.

### Sample size

The sample size calculation was based on the desired accuracy of the limits of agreement.<sup>23</sup> Assuming an accuracy of  $\pm 0.075$  °C for the 95% CIs of the limits of agreement (and the standard deviation of the agreement between temperatures measured by non-contact and electronic axillary thermometers would be 0.5 °C, as documented in previous studies<sup>9–12,14</sup>), a minimum sample size of 533 participants would have been required. We revised this sample size on 14 May 2018. Based on the data already available, the standard deviation for our primary analysis (Thermofocus non-contact thermometer vs. electronic axillary thermometer) was 0.65 °C. Using this standard deviation, the original sample size of 533 children would give us 0.10 °C accuracy for each limit of agreement (rather than 0.075 °C as anticipated). A reduced sample size of 400 children would give us 0.11 °C accuracy. Considering that thermometers measure temperatures only to the closest 0.1 °C, we felt that this would be sufficient as the rounding would make the two estimates equivalent. This change was discussed and approved by the funder. Secondary outcomes of agreement between the other thermometer types have been estimated with the same precision.

### Protocol changes

Other than the change in sample size, there have been no changes to the protocol after the start of patient recruitment.

### Analyses

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.

Our primary outcome is the agreement between the Thermofocus NCIT and the electronic axillary thermometer. Analyses of agreement were conducted based on Bland-Altman plots, which provide an indication of bias and limits of agreement between the measurements. Exact 95% CIs around this estimate have been calculated.

Diagnostic accuracy for detecting fever (temperature of  $\geq 38$  °C measured by the electronic axillary thermometer) was analysed by calculating sensitivity, specificity, predictive values and likelihood ratios, with 95% CIs. Other analyses, such as failure rates, are reported as proportions.

The scores on the visual analogue scale and the Patient Discomfort Scale have been analysed using non-parametric techniques resulting in median acceptability [and interquartile ranges (IQRs)] for each thermometer. Thermometer preference was analysed via chi-squared goodness-of-fit testing.

## Qualitative data collection and analysis

### Recruitment

Parents were purposively sampled from those consenting to contact to achieve maximum variation in gender, age of parent, age of child, ethnicity and number of other children in the household. Recruitment continued until the research team agreed that data saturation had been achieved and sufficient explanation for the categories generated was reached.

### Data collection

Interviews were conducted by telephone ( $n = 20$ ) or face to face ( $n = 1$ ), according to participant preference, from May 2017 to June 2018. All participants gave written or recorded verbal informed consent prior to the interview.

Interviews were conducted separately by two researchers trained in qualitative methodology; Elizabeth Morris (a female clinical researcher and salaried GP) and Fatene Abakar Ismail (a female research assistant). Continuity was ensured by regular discussion between researchers, and review of transcripts and topic guides by the team (EM, FAI, GHa and MG).

Interviews were semistructured and followed a flexible topic guide developed by the research team and patient and public involvement (PPI) panel. Content of the topic guide was informed by the primary aims of the study, the existing literature on parents' experiences of childhood fever<sup>24–28</sup> and the expertise of the research team. The topic guide was reviewed iteratively by the research team and evolved in response to emerging themes and ongoing PPI. In particular, the attribute of 'accuracy' emerged as an early theme and was incorporated into the topic guide to allow further detailed exploration. Interviews lasted 10–15 minutes, on average, and were audio-recorded, transcribed verbatim by a transcription company and checked against the original recording by the research team.

The study was approved by the South Central – Berkshire Research Ethics Committee (reference 17/SC/0068).

### Data analysis

Data analysis followed a thematic approach. Following transcription, the researchers undertook 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 (EM and MG), with the assistance of NVivo (version 11; QSR International, Warrington, UK) qualitative data analysis software. Interpretation of the codes and themes was developed in discussion with the wider research team.



## Chapter 3 Results

We recruited 401 children with a median age of 1.62 years (IQR 0.79–3.39 years), 203 (50.62%) of whom were boys. Most children were of white British ethnicity (69.83%). Approximately 30% of children were feverish at the time of inclusion and 33% were given fever medication in the previous 6 hours. Patient characteristics are listed in *Table 1*.

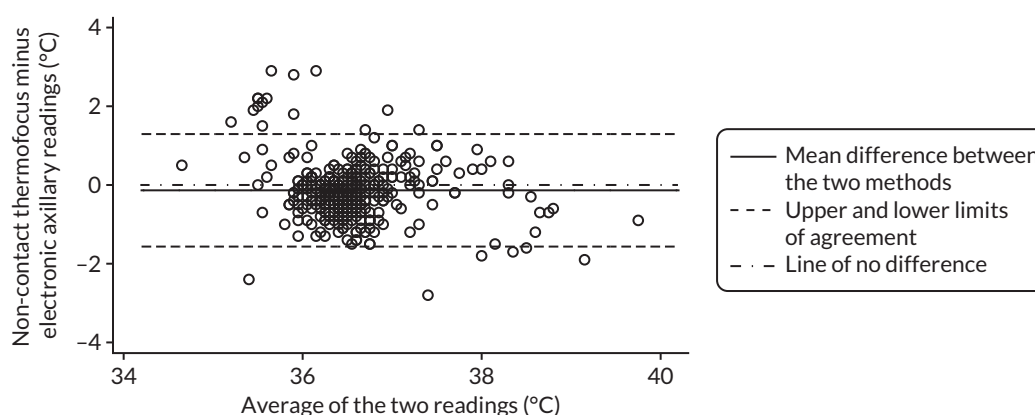
There were 396 temperature readings with the Thermofocus non-contact thermometer (first measurement), 399 for the Firhealth non-contact thermometer (first measurement), 376 for the electronic axillary thermometer and 390 for the tympanic thermometer. Second measurements with the Thermofocus and Firhealth non-contact thermometers resulted in 395 and 397 readings, respectively.

The Bland–Altman plot for our primary outcome, which is the agreement between the Thermofocus non-contact thermometer (first measurement) and the electronic axillary thermometer, is presented in *Figure 1*. The mean difference between axillary and non-contact was  $-0.14^{\circ}\text{C}$  (95% CI  $-0.21$  to  $-0.06^{\circ}\text{C}$ ), with the lower limit of agreement being  $-1.57^{\circ}\text{C}$  (95% CI  $-1.69$  to  $-1.44^{\circ}\text{C}$ ) and the upper limit being  $1.29^{\circ}\text{C}$  (95% CI  $1.16$  to  $1.42^{\circ}\text{C}$ ). This means that in 95 out of 100 cases we would expect the difference between

TABLE 1 Participant characteristics

Variable	Summary statistic
Age (years), median (IQR)	1.62 (0.79–3.39)
Gender (boys), <i>n</i> (%)	203 (50.62)
Ethnicity, <i>n</i> (%)	
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, <i>n</i> (%)	119 (29.75)
Fever medication in past 6 hours, <i>n</i> (%)	134 (33.50)
Fever duration (days), median (IQR)	1.5 (0.5–3)
Illness duration (days), median (IQR)	3 (2–7)
Setting, <i>n</i> (%)	
Out-of-hours primary care	34 (8.48)
In-hours primary care	367 (91.52)

Reproduced with permission from Hayward *et al.*<sup>15</sup>



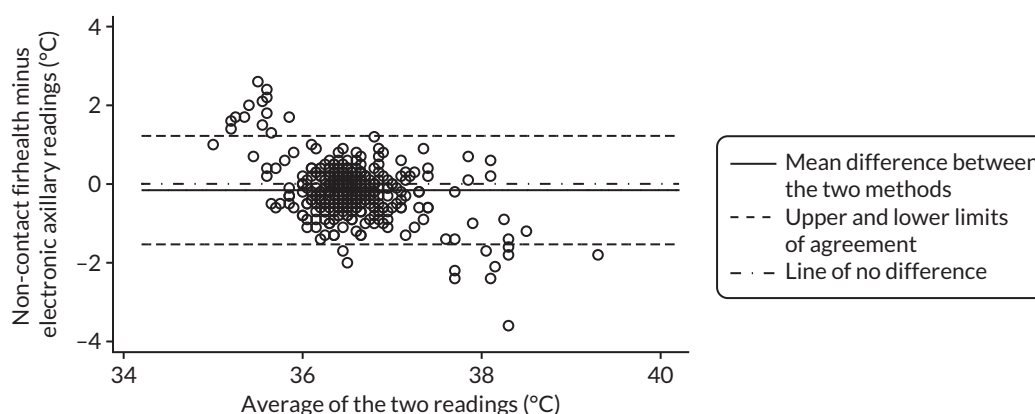
**FIGURE 1** Bland-Altman plot for agreement between the Thermofocus non-contact thermometer and the electronic axillary thermometer. Reproduced with permission from Hayward *et al.*<sup>15</sup>

the axillary and the non-contact thermometer to range between 1.57 °C lower and 1.29 °C higher than the average of the non-contact and axillary measurement of temperature.

The mean difference between the second non-contact thermometer (Firhealth, first measurement) 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) (Figure 2).

All method comparison results, including non-contact versus tympanic and non-contact first and second measurements, are summarised in Table 2. Accompanying plots are available in Appendix 2. When assessed against the tympanic thermometer, agreement was similarly dependent on average temperature.

We calculated accuracy of non-contact thermometers for diagnosing fever, defined as a temperature of  $\geq 38$  °C by electronic axillary measurement. Prevalence of fever thus defined was 4.26% (95% CI 2.45% to 6.82%). Sensitivity and specificity for the Thermofocus non-contact thermometer were 66.7% (95% CI 38.4% to 88.2%) and 98.0% (95% CI 96.0% to 99.2%), respectively. For the Firhealth thermometer, sensitivity was 12.5% (95% CI 1.6% to 38.3%) and specificity was 99.4% (95% CI 98.0% to 99.9%). The sensitivity of the other standard thermometer, the tympanic thermometer, was moderate (62.5%, 95% CI 35.4% to 84.8%). Table 3 shows further details.



**FIGURE 2** Bland-Altman plot for agreement between the Firhealth non-contact thermometer and the electronic axillary thermometer. Reproduced with permission from Hayward *et al.*<sup>15</sup>

TABLE 2 Method agreement results

Comparison	Mean difference and 95% CI (°C)	Lower limit of agreement and 95% CI (°C)	Upper limit of agreement and 95% CI (°C)
Thermofocus NCIT minus electronic axillary, <i>n</i> = 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, <i>n</i> = 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, <i>n</i> = 384	-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, <i>n</i> = 387	-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, <i>n</i> = 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 infrared tympanic, <i>n</i> = 365	0.06 (-0.02 to 0.14)	-1.49 (-1.63 to -1.34)	1.61 (1.47 to 1.75)
Thermofocus NCIT first minus second reading, <i>n</i> = 395	-0.04 (-0.07 to -0.01)	-0.56 (-0.60 to -0.51)	0.47 (0.43 to 0.52)
Firhealth non-contact first minus second reading, <i>n</i> = 396	0.01 (-0.02 to 0.04)	-0.60 (-0.65 to -0.54)	0.61 (0.56 to 0.67)
Reproduced with permission from Hayward <i>et al.</i> <sup>15</sup>			

TABLE 3 Diagnostic accuracy for fever defined as  $\geq 38^{\circ}\text{C}$  by the electronic axillary thermometer

Comparison	Thermofocus	Firhealth	Tympanic
Sensitivity, % (95% CI)	66.7 (38.4 to 88.2)	12.5 (1.6 to 38.3)	62.5 (35.4 to 84.8)
Specificity, % (95% CI)	98.0 (96.0 to 99.2)	99.4 (98.0 to 99.9)	96.0 (93.4 to 97.8)
Positive predictive value, % (95% CI)	58.8 (32.9 to 81.6)	50.0 (6.8 to 93.2)	41.7 (22.1 to 63.4)
Negative predictive value, % (95% CI)	98.6 (96.7 to 99.5)	96.2 (93.7 to 97.9)	98.2 (96.2 to 99.4)
Likelihood ratio + (95% CI)	33.9 (15.0 to 76.7)	22.4 (3.4 to 148.8)	15.6 (8.2 to 29.5)
Likelihood ratio - (95% CI)	0.34 (0.17 to 0.70)	0.88 (0.73 to 1.06)	0.39 (0.21 to 0.74)
Absolute numbers (true positive, false positive, false negative, true negative)	10, 5, 7, 349	2, 14, 2, 356	10, 6, 14, 335
Reproduced with permission from Hayward <i>et al.</i> <sup>15</sup>			

The number of attempts for each thermometer that were required to obtain a valid reading and the number of failed readings are listed in *Table 4*.

## Acceptability of the thermometers

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

Overall, most children were rated by their parents as relaxed during temperature measurements with each thermometer. *Figure 3* displays the distribution of the Patient Discomfort Scale for all four thermometers for the children in whom this was assessed (axillary, *n* = 297; tympanic, *n* = 305; Thermofocus, *n* = 304; Firhealth, *n* = 305). The median for each thermometer was 2, although the IQR was slightly larger for the axillary thermometer than for the other thermometers: Thermofocus median 2 (IQR 2–2), Firhealth median 2 (IQR 2–2), electronic axillary median 2 (IQR 2–3) and infrared tympanic median 2 (IQR 2–2).

## RESULTS

TABLE 4 Number of attempts for each thermometer

Comparison	Thermofocus non-contact, n (%)	Firhealth non-contact, 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)
<b>No reading</b>				
Technical error: thermometer not activating	3 (0.8)	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 co-operation 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 (child aged < 4 weeks)	0 (0.0)	0 (0.0)	0 (0.0)	5 (1.2)
Total	401	401	401	401
Reproduced with permission from Hayward <i>et al.</i> <sup>15</sup>				

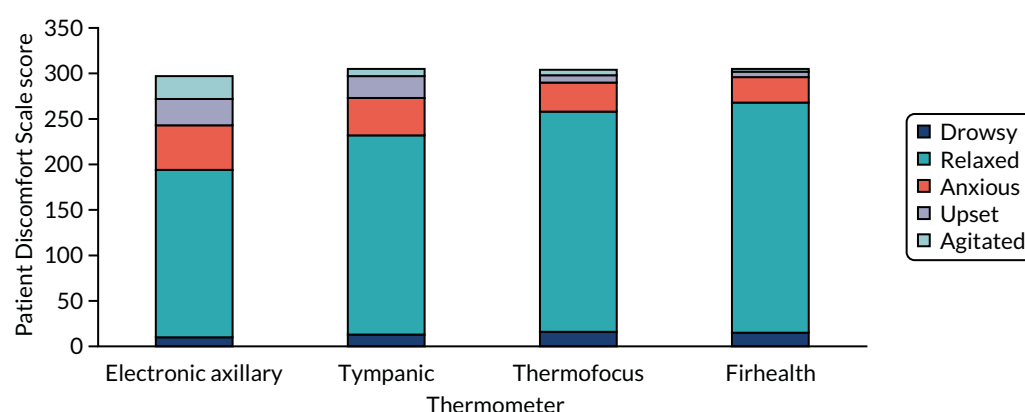


FIGURE 3 Results of the Patient Discomfort Scale.

Parental acceptability as assessed by a visual analogue scale was highest for the Firhealth non-contact thermometer (median 9 cm, IQR 7.6–9.5 cm), followed by the Thermofocus non-contact thermometer (median 8.5 cm, IQR 6.9–9.4 cm), the tympanic thermometer (median 7.6 cm, IQR 5.5–9 cm) and, finally, the electronic axillary thermometer (median 5 cm, IQR 2.1–7.6 cm).

When asked to rank the four thermometers by preference, with 1 indicating most preferred and 4 indicating least preferred, most parents preferred the Firhealth non-contact thermometer (median 2, IQR 1–2), followed by the Thermofocus (median 2, IQR 1–3), the tympanic thermometer (median 3, IQR 2–3) and, finally, the electronic axillary thermometer (median 4, IQR 3–4). However, these differences were not statistically significant (Friedman  $p = 1.00$ ).

## Qualitative interviews

A total of 65% of the participants in the METRIC study consented to be approached for the nested qualitative interviews. Interviews were conducted with 21 parents who had been purposively sampled to achieve a maximum variation sample, and recruitment continued until data saturation was achieved. The characteristics of 21 parents who participated in the interviews are described in the *Appendix 1, Table 6*.



The main themes concerning the individual device attributes are summarised in *Table 5*, and described below.

### **Electronic axillary thermometer**

Parents' experiences with the axillary thermometers were described using negative language in almost all cases, with only two participants – both older parents – expressing neutral feelings towards them ('He was fine, he was quite happy with those'; i11, mother, white British, aged 31–40 years, one child).

A commonly described concern was the child's perceived discomfort during use of the axillary thermometer, with more than half of respondents describing their child as appearing 'uncomfortable' during its use. Some participants felt this to be a more physical discomfort for the child, whereas, for others, the appearance of the device contributed to the child's negative experience:

*He freaked out at that one ... Because he thought it was a needle ... so he thought he was going to have an injection.*

*i8, mother, white British, aged 31–40 years, two children*

Barriers to the convenience and practicality of the procedure required to measure the temperature with this device were also frequently raised by parents, with some feeling that having to undress the child to use the axillary thermometer was impractical for the parent and contributed to the child's discomfort when they were already feeling unwell:

*The armpit kind of seemed impractical ... you have to undress your kid to jam a big metal stick under his arm.*

*i1, father, white British, two children*

However, despite these experiences, several parents still expressed trust that if it was the medically recommended thermometer to use, they would be willing to persevere because the inconvenience or discomfort was temporary:

*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, makes no, there was no harm.*

*i1, father, white British, two children*

**TABLE 5** Summary of main themes of the qualitative interviews

Thermometer	Device attributes	Comments
Axillary	Discomfort	Predominantly negative terminology used to describe the parental and child experience
	Impracticality	
	Appearance	
Tympanic	Familiarity	More neutral vocabulary and terminology used to describe the parental and child experience
	Neutral convenience and comfort	
NCIT	Convenience	Most positive terminology used regarding ease and practicality of use; however, reassurance needed regarding the safety of newer technology
	Practicality	
	Non-invasive	
	Safety	

### **Infrared tympanic thermometer**

The parent descriptions of their child's experience of temperature measurement with tympanic thermometers was more neutral than with the axillary thermometers, demonstrating more acceptance of the milder discomfort or displeasure displayed by their children in this context:

*Yes she was used to that, she said it would be like a [um] that facial expression like very curious, what is that. But not really rigid not really in distress, no not at all.*

*i13, mother, aged > 40 years, one child*

Several parents related this fact to the thermometer being the one their child was used to from being used in the home environment. However, a commonly emerging theme regarding the tympanic thermometer was parent concern about user-introduced variability, and that their skills in using it in the home environment might be affecting the results:

*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, white British, aged 31–40 years, one child*

*You feel like there are so many ways that you can get it into their ear and you don't know whether it needs to be really deep and like properly in or, you know, it comes against the part of the ear you don't know whether it's been an accurate reading or if you don't have any background, medical background you don't really know which, you know, if there are many ways in which it goes in is more effective or what. When you're stressed and worried there's lots of various things to worry about.*

*i6, mother, white British, aged 31–40 years, one child*

### **Non-contact infrared thermometers**

In comparison with the tympanic and axillary thermometers, the Thermofocus infrared thermometer was felt to provide feedback on 'correct' usage with its visible light feature, and on interpretation of the results with symbols, which were described as positive and reassuring components by several participants:

*With the blue one there was a light, I thought that was quite helpful so you knew [um] 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, white British, aged 31–40 years, one child*

*The blue one had like a smiley face didn't it so just kind of symbols rather than numbers is that right? I think I much preferred that because even though I would count myself as a fairly [um] knowledgeable person, when she's unwell I'm a bit stressed I can't then remember what counts as a temperature.*

*i6, mother, white British, aged 31–40 years, one child*

Practicality and convenience of the procedure for parents and children emerged again as a theme, but was felt here to be a positive feature of the NCITs. The fact that the child might not even be aware of the temperature measurement being taken improved the experience from both the perspective of the parent and their perception of the child's experience:

*For ease, the forehead one where you don't even touch them is an easy one isn't it because ... like you could do it while they're asleep, you know, from a practical point of view that would probably be the easiest. Like I said it's less intrusive [um] I think it's a bit more flexible because you can do it when they're asleep.*

*i1, father, white British, two children*

*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, white British, one child, with history of fever this episode*

However, parents were, in general, less aware of this newer type of non-contact thermometer and expressed a wish for reassurance or guidance that they would be reliable, given their significant differences when compared with more traditional methods of temperature measurement:

*I think it's mainly because you're kind of used to the [um] the touch the temperature when you get, when it's touched the body and you know, like that physical contact for 1 or 2 minutes, so if you [um] kind of or 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*



## Chapter 4 Patient and public involvement

The established PPI panel for the NIHR Diagnostic Evidence Cooperative Oxford agreed to adopt this study in to its portfolio of projects. We have discussed this project with our panel (which includes two mothers of children aged < 5 years) from inception until reporting of the final results.

The panel had access to the Department of Primary Care Health Sciences' information technology and library facilities, and training opportunities were identified and offered to the group. They were reimbursed for their time, travel and childcare expenses.

The PPI panel provided feedback on the relevance of the topic itself, provided feedback on the study materials (particularly the patient information leaflets and acceptability ratings by children and parents), commented on the topic guide for the qualitative interviews with the parents and provided feedback on our sample size revision. In addition, we sought their advice on acceptable limits of agreement between thermometry devices, and discussed our emerging themes on the acceptability of the thermometer types from our qualitative interviews. We also asked for advice on how to disseminate our findings to a lay audience.

We discussed the initial study proposal with the NIHR Diagnostic Evidence Cooperative Oxford PPI panel and with a team of nursery nurses. They saw clear advantages to a non-contact thermometry device, particularly in eliminating the need for expensive and inconvenient disposable covers. One panel member reported that she owned a NCIT and doubted its accuracy; it provided lower readings than other thermometers. They felt that the study question was timely and important for parents as well as GPs. The issue of reproducibility of readings was raised and, in response, we have expanded our objectives to include an evaluation of NCIT reproducibility. They suggested that the cost of the Thermofocus could be prohibitive to adoption into primary care and nursery environments and were enthusiastic about the idea of including a cheaper comparison brand. The panel suggested modifications to the study flow to allow parents to remain in the GP waiting room for the study, or to participate after their GP appointment should they find this more convenient. They reviewed and clarified our *Plain English summary*.

The proposed change in sample size was discussed, in particular, with the two parents with young children who have commented on the study previously. Both parents felt that the change would still leave us with an acceptable amount of precision. One had tried all three types on her toddler and commented that the intermeasurement variability of one particular thermometer was far higher than this for a single thermometer (e.g. when measured in both ears by a tympanic thermometer). All acknowledged the need to balance finances and time with precision and thought that this was an acceptable compromise. One felt that it was important that we were transparent about this change and the rationale in our write-up. One felt that 400 was still a very large number of children to recruit to a study of this type.

We discussed the content of the topic guide with the PPI panel, and initial results from its pilot use; we discussed how to add the emerging theme of accuracy into the topic guide, and the panel made particular recommendations to explore parents' feelings about cost of thermometers (and what might be acceptable to parents) in our qualitative interviews.

The PPI panel has also reviewed our final results and were very enthusiastic about them. The PPI panel felt that an article for the general media should definitely be written. The PPI panel commented that they and most of their friends owned any combination of the thermometers used in the study. The PPI panel commented that given the purchase price of these thermometers, it would be great if people were informed about their performance so that they could consider whether or not/what to purchase.

The PPI panel also suggested the potential to distribute leaflets (giving the study outcomes) through nurseries, schools, posters in GP surgeries (flash-up screens in waiting rooms) and even letters to parents at schools, although it realised that this may be quite ambitious and possibly not feasible.

# Chapter 5 Discussion

## Summary of the main findings

When comparing non-contact thermometers with currently recommended thermometers in young children (with a maximum age of 5 years), the mean difference between thermometers was not very large but the upper and lower limits of agreement were  $> 1^{\circ}\text{C}$  for the comparison with both the electronic axillary and the infrared tympanic thermometer. As a high temperature is considered to be a red flag for serious infections, this means that in a substantial number of children the differences between temperature readings could affect clinical decision-making.<sup>4</sup> The limits of agreement also surpass the commonly accepted limits of agreement of  $0.5^{\circ}\text{C}$ .<sup>29</sup>

This matters because high temperatures, especially of  $\geq 40^{\circ}\text{C}$ , are associated with an increased risk of serious infection that can have very adverse consequences if left undetected and untreated.<sup>30</sup> In our study, two children with a temperature reading of at least  $40^{\circ}\text{C}$  with the axillary thermometer had temperature readings of  $38.2$  and  $39.3^{\circ}\text{C}$  with the Thermofocus non-contact thermometer, and  $36.5$  and  $38.4^{\circ}\text{C}$  with the Firhealth non-contact thermometer. As is the nature of method comparison studies where there are no assumptions on which of the methods is the reference standard, it is impossible to say whether this pattern is caused by the non-contact thermometers or the axillary or tympanic thermometers.

We found that reproducibility of measurements with the non-contact thermometers was good, with very small differences between the first and the second reading of each non-contact thermometer. In addition, there were very small differences between both non-contact thermometers with the upper and lower limits of agreement being  $< 1^{\circ}\text{C}$ . Fewer attempts and fewer failures were reported for the non-contact thermometers than for the axillary or tympanic thermometers.

Parental satisfaction and child discomfort appear to be better for the non-contact thermometers but the majority rated all devices as acceptable. Based on their ranking, discomfort, pain score and acceptability, parents and children least preferred the axillary thermometer. This type of thermometer also resulted in the most failed readings. The Firhealth non-contact thermometer was most liked by parents and associated with the lowest discomfort for children. In the interviews, parents confirmed that they preferred the practicality and comfort of non-contact thermometers. In addition, they mostly used negative language when talking about the axillary thermometer, but they also indicated that they would still use it if this type of thermometer was recommended by clinicians.

## Strengths and limitations

We recruited 401 children who attended GP settings with an acute illness, making this one of the larger studies to assess peripheral thermometers in routine clinical settings. We also employed a mixed-methods approach, combining and integrating quantitative and qualitative analyses. We randomised the order in which thermometers were used and compared two new types with two currently recommended types.

There are some limitations to our study. First, uncertainty exists over the accuracy of the currently recommended thermometers that we included as comparisons (i.e. the electronic axillary thermometer and infrared tympanic thermometer). In a recent systematic review,<sup>16</sup> it was found that neither method met acceptable limits of agreement compared with core temperature measurements. This conclusion is reinforced by our comparison of the axillary and tympanic thermometers, which also showed wide limits of agreement of  $> 1.5^{\circ}\text{C}$ , which is well beyond the  $0.5^{\circ}\text{C}$  threshold of acceptability. This, in turn, makes it challenging to interpret our findings. Second, we recruited fewer children who were febrile at

the time of measurement than we had anticipated, meaning that the CIs around our estimate of sensitivity are wider than we would have liked. In an attempt to avoid this, our research team approached consecutive children arriving at GP waiting rooms, but the parents who agreed to participate may have been those with children who they felt were less unwell, a component of which could have been fever. 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 what would be expected among practice staff or parents. However, all of the methods were simple to use and could be used frequently in most primary care settings, meaning that expertise would be likely to rapidly develop.

We encountered some difficulties in recruiting a sufficient number of patients in the time provided, which resulted in a loss of power. Recruitment problems were partly due to the complexity of the protocol, which required research assistants to go into the practices for recruitment rather than relying on opportunistic recruitment by practice staff. Research assistant time was further restrained by our study budget, meaning that hours of recruitment were ultimately limited. In addition, as children could not be entered in the study more than once, we reached saturation in practices that had collaborated in the study from the start, meaning that new practices had to be recruited that were increasingly further away.

## Comparison with existing literature

We found a mean difference of  $-0.14$  and  $-0.16$  °C between NCITs and electronic axillary thermometers, which is within the range demonstrated by other studies comparing NCITs with electronic axillary, infrared tympanic and rectal thermometers. A study in 50 hospitalised children up to 14 years of age found a difference of  $-0.38$  °C between electronic axillary and non-contact thermometers, with limits of agreement of  $-1.47$  to  $0.70$  °C.<sup>18</sup> Another study in hospitalised children aged up to 17 years ( $n = 294$ ) compared a tympanic thermometer (AccuSystem Genius2; Covidien, Dublin, Republic of Ireland), a NCIT (ThermoFlash Contactless; Visiomed Group SA, Paris, France) and a temporal artery scan thermometer (Exergen; Exergen, Watertown, MA, USA) with an electronic thermometer used rectally (Filac 3000; Covidien, Dublin, Republic of Ireland), and found that the mean difference between the non-contact and rectal thermometer was  $0.34$  °C (limits of agreement  $-0.92$  to  $1.60$  °C).<sup>19</sup> Finally, Teran *et al.*<sup>10</sup> compared a non-contact thermometer (Thermofocus) and a temporal artery thermometer (Exergen), which is also infrared but requires minimal skin contact, with a mercury-in-glass thermometer used rectally in 500 children aged up to 48 months, half of whom were hospitalised and the other half presenting at the emergency room. They found a mean difference between the rectal reading and the non-contact reading of  $0.029$  °C (limits of agreement  $-0.50$  to  $0.58$  °C). The study by Chiappini *et al.*<sup>9</sup> showed greater agreement between the Thermofocus and an axillary measurement, but they used a mercury-in-glass thermometer for the axillary temperature reading. They also used the average of two axillary temperature readings and three non-contact temperature readings.

The underestimation of fever was also reported in three studies: one comparing the Thermofocus NCIT with rectal thermometry in 200 children attending the emergency department,<sup>20</sup> another of 294 paediatric inpatients comparing the ThermoFlash NCIT with rectal thermometry as described above<sup>19</sup> and another of 119 healthy newborns comparing the Thermofocus with an electronic axillary thermometer and tympanic thermometer.<sup>31</sup> This trend was also noted in a systematic review of temporal artery thermometers, which also use infrared technology, compared with other methods.<sup>29</sup>

## Implications for research and practice

In children aged  $\leq 5$  years, the temperature measured by non-contact thermometers can differ from measurements with an electronic axillary or infrared tympanic thermometer by  $> 1$  °C, and the agreement appears to be affected by the average temperature. Given the uncertainty over the accuracy of currently recommended thermometers, it is hard to draw firm conclusions about the likely



impact on practice if non-contact thermometers were introduced as standard care. If we assume that electronic axillary thermometry is an accurate reference standard, then the moderate agreement between non-contact thermometers and an electronic axillary reading, and low sensitivity for fever mean that primary care clinicians should be cautious in using this technology.

However, we also found wide limits of agreement between electronic axillary and infrared tympanic thermometers, both of which are advocated in guidance. Furthermore, parents were more positive about the benefits of non-contact thermometers than about the electronic axillary and infrared tympanic thermometers, as evidenced by both rating scales and qualitative interviews, and there was good agreement between repeated measurements, which suggested that multiple attempts would not be required.

Therefore, clinicians need to be cautious about the accuracy of any peripheral thermometry approach and ensure that the management decisions they make are using these data as part of a holistic assessment. There is clear potential for technological innovation in this field to develop novel, more accurate methods of peripheral thermometry to support clinical decision-making. Once clinical evidence is more firmly established, an economic evaluation balancing clinical effectiveness with costs could be useful.



# Acknowledgements

## Contributions of authors

**Professor Ann Van den Bruel** (Associate Professor and GP) and **Professor Gail Hayward** (<https://orcid.org/0000-0003-0852-627X>) (Associate Professor and GP) conceived the study, provided oversight of study processes and drafted the manuscript. **Ann Van den Bruel** was responsible for submitting the final report.

**Professor Jan Verbakel** (Assistant Professor and GP) contributed to protocol development, interpretation of the findings and drafting of the manuscript.

**Dr Kay Wang** (<https://orcid.org/0000-0002-7195-1730>) (NIHR Postdoctoral Fellow and GP) contributed to protocol development, interpretation of the findings and drafting of the manuscript.

**Dr Susannah Fleming** (<https://orcid.org/0000-0001-7205-2051>) (Quantitative Researcher) contributed to protocol development, in particular the sample size calculations and analyses plan, interpretation of the findings and drafting of the manuscript.

**Dr Gea Holtman** (<https://orcid.org/0000-0001-6579-767X>) (Quantitative Researcher) contributed to data collection, interpretation of the findings and drafting of the manuscript.

**Dr Margaret Glogowska** (<https://orcid.org/0000-0001-8029-1052>) (Senior Health Services Researcher) and **Elizabeth Morris** (<https://orcid.org/0000-0002-9913-6041>) (NIHR In-practice Fellow and GP) contributed to protocol development, in particular the nested qualitative study, analysed the qualitative data, interpreted the findings and drafted the manuscript.

**Mr George Edwards** (<https://orcid.org/0000-0002-0048-4178>) (Research Assistant) contributed to data collection, data management, analyses and drafting of the manuscript.

**Dr Fatene Abakar Ismail** (<https://orcid.org/0000-0002-1463-7980>) (Trial Manager) contributed to protocol development, data collection and drafting of the manuscript.

**Ms Kathryn Curtis** (<https://orcid.org/0000-0001-5367-9182>) (Research Assistant) contributed to data collection, interpretation of the findings and drafting of the manuscript.

**Dr James Goetz** (<https://orcid.org/0000-0003-2439-8116>), **Dr Grace Barnes** (<https://orcid.org/0000-0003-4296-1689>), **Dr Ralitsa Slivkova** (<https://orcid.org/0000-0002-2810-2266>), **Miss Charlotte Nesbitt** (<https://orcid.org/0000-0002-2873-153X>), **Dr Suhail Aslam**, (<https://orcid.org/0000-0003-4205-814X>), **Dr Ealish Swift** (<https://orcid.org/0000-0001-6884-4881>) and **Dr Harriet Williams** (<https://orcid.org/0000-0002-8009-9236>) (medical students) contributed to data collection, interpretation of the findings and drafting of the manuscript.

## Publication

Hayward G, Verbakel JY, Ismail FA, Edwards G, Wang K, Fleming S, *et al.* Non-contact infrared versus axillary and tympanic thermometers in children attending primary care: a mixed-methods study of accuracy and acceptability. *Br J Gen Pract* 2020;**70**:e236–44. [Corrigendum published in *Br J Gen Pract* 2020;**70**:486.]

## **Data-sharing statement**

All data requests should be submitted to the corresponding author for consideration. Access to anonymised data may be granted following review.

## **Patient data**

This work uses data provided by patients and collected by the NHS as part of their care and support. Using patient data is vital to improve health and care for everyone. There is huge potential to make better use of information from people's patient records, to understand more about disease, develop new treatments, monitor safety, and plan NHS services. Patient data should be kept safe and secure, to protect everyone's privacy, and it's important that there are safeguards to make sure that it is stored and used responsibly. Everyone should be able to find out about how patient data are used. #datasaveslives You can find out more about the background to this citation here: <https://understandingpatientdata.org.uk/data-citation>.

# References

1. Hay AD, Heron J, Ness A, ALSPAC study team. 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**:367–74. <https://doi.org/10.1093/fampra/ctmi035>
2. Fleming DM, Smith GE, Charlton JR, Charlton J, Nicoll A. Impact of infections on primary care – greater than expected. *Commun Dis Public Health* 2002;**5**:7–12.
3. Armon K, MacFaul R, Hemingway P, Werneke U, Stephenson T. The impact of presenting problem based guidelines for children with medical problems in an accident and emergency department. *Arch Dis Child* 2004;**89**:159–64. <https://doi.org/10.1136/adc.2002.024406>
4. National Institute for Health and Care Excellence. *Fever in Under 5s: Assessment and Initial Management*. NICE Guidelines [CG160]. URL: [www.nice.org.uk/guidance/cg160](http://www.nice.org.uk/guidance/cg160) (accessed 26 June 2019).
5. Omron®. *Omron® 20-second Flexible Digital Thermometer Model # MC-206 Instructions*. URL: [www.manualslib.com/manual/504879/Omron-Mc-206.html](http://www.manualslib.com/manual/504879/Omron-Mc-206.html) (accessed 26 June 2019).
6. Craig JV, Lancaster GA, Taylor S, Williamson PR, Smyth RL. Infrared ear thermometry compared with rectal thermometry in children: a systematic review. *Lancet* 2002;**360**:603–9. [https://doi.org/10.1016/S0140-6736\(02\)09783-0](https://doi.org/10.1016/S0140-6736(02)09783-0)
7. Wang K, Gill P, Wolstenholme J, Price CP, Heneghan C, Thompson M, Plüddemann A. Non-contact infrared thermometers for measuring temperature in children: primary care diagnostic technology update. *Br J Gen Pract* 2014;**64**:e681–3. <https://doi.org/10.3399/bjgp14X682045>
8. National Institute for Health Research. *Horizon Scanning Reports*. URL: [www.community.healthcare.mic.nihr.ac.uk/reports-and-resources/horizon-scanning-reports](http://www.community.healthcare.mic.nihr.ac.uk/reports-and-resources/horizon-scanning-reports) (accessed December 2019).
9. Chiappini E, Sollai S, Longhi R, Morandini L, Laghi A, Osio CE, *et al.* Performance of non-contact infrared thermometer for detecting febrile children in hospital and ambulatory settings. *J Clin Nurs* 2011;**20**:1311–18. <https://doi.org/10.1111/j.1365-2702.2010.03565.x>
10. Teran CG, Torrez-Llanos J, Teran-Miranda TE, Balderrama C, Shah NS, Villarroel P. Clinical accuracy of a non-contact infrared skin thermometer in paediatric practice. *Child Care Health Dev* 2012;**38**:471–6. <https://doi.org/10.1111/j.1365-2214.2011.01264.x>
11. Fortuna EL, Carney MM, Macy M, Stanley RM, Younger JG, Bradin SA. Accuracy of non-contact infrared thermometry versus rectal thermometry in young children evaluated in the emergency department for fever. *J Emerg Nurs* 2010;**36**:101–4. <https://doi.org/10.1016/j.jen.2009.07.017>
12. Selent MU, Molinari NM, Baxter A, Nguyen AV, Siegelson H, Brown CM, *et al.* Mass screening for fever in children: a comparison of 3 infrared thermal detection systems. *Pediatr Emerg Care* 2013;**29**:305–13. <https://doi.org/10.1097/PEC.0b013e3182854465>
13. 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**:327–36.
14. Ng KG, Wong ST, Lim SM, Goh Z. Evaluation of the Cadi ThermoSENSOR wireless skin-contact thermometer against ear and axillary temperatures in children. *J Pediatr Nurs* 2010;**25**:176–86. <https://doi.org/10.1016/j.pedn.2008.12.002>
15. Hayward G, Verbakel JY, Ismail FA, Edwards G, Wang K, Fleming S, *et al.* Non-contact infrared versus axillary and tympanic thermometers in children attending primary care: a mixed-methods study of accuracy and acceptability. *Br J Gen Pract* 2020;**70**:e236–44. [Corrigendum published in *Br J Gen Pract* 2020;**70**:486.]

16. Niven DJ, Gaudet JE, Laupland KB, Mrklas KJ, Roberts DJ, Stelfox HT. Accuracy of peripheral thermometers for estimating temperature: a systematic review and meta-analysis. *Ann Intern Med* 2015;**163**:768–77. <https://doi.org/10.7326/M15-1150>
17. 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**:1174–8. <https://doi.org/10.1136/bmj.320.7243.1174>
18. Apa H, Gözmen S, Bayram N, Çatkoğlu A, Devrim F, Karaarslan U, *et al.* Clinical accuracy of tympanic thermometer and noncontact infrared skin thermometer in pediatric practice: an alternative for axillary digital thermometer. *Pediatr Emerg Care* 2013;**29**:992–7. <https://doi.org/10.1097/PEC.0b013e3182a2d419>
19. Allegaert K, Casteels K, van Gorp I, Bogaert G. Tympanic, infrared skin, and temporal artery scan thermometers compared with rectal measurement in children: a real-life assessment. *Curr Ther Res Clin Exp* 2014;**76**:34–8. <https://doi.org/10.1016/j.curtheres.2013.11.005>
20. Paes BF, Vermeulen K, Brohet RM, van der Ploeg T, de Winter JP. Accuracy of tympanic and infrared skin thermometers in children. *Arch Dis Child* 2010;**95**:974–8. <https://doi.org/10.1136/adc.2010.185801>
21. Greenes DS, Fleisher GR. Accuracy of a noninvasive temporal artery thermometer for use in infants. *Arch Pediatr Adolesc Med* 2001;**155**:376–81. <https://doi.org/10.1001/archpedi.155.3.376>
22. Hockenberry M, Wilson D, Winkelstein M. *Wong's Essentials of Pediatric Nursing*. 7th edn. St. Louis, MO: Mosby; 2005.
23. Bland JM, Altman DG. Agreed statistics: measurement method comparison. *Anesthesiology* 2012;**116**:182–5. <https://doi.org/10.1097/ALN.0b013e31823d7784>
24. Kelly M, McCarthy S, O'Sullivan R, Shiely F, Larkin P, Brenner M, Sahm LJ. Drivers for inappropriate fever management in children: a systematic review. *Int J Clin Pharm* 2016;**38**:761–70. <https://doi.org/10.1007/s11096-016-0333-2>
25. Kelly M, Sahm LJ, Shiely F, O'Sullivan R, McGillicuddy A, McCarthy S. Parental knowledge, attitudes and beliefs regarding fever in children: an interview study. *BMC Public Health* 2016;**16**:540. <https://doi.org/10.1186/s12889-016-3224-5>
26. de Bont EG, Loonen N, Hendrix DA, Lepot JM, Dinant GJ, Cals JW. Childhood fever: a qualitative study on parents' expectations and experiences during general practice out-of-hours care consultations. *BMC Fam Pract* 2015;**16**:131. <https://doi.org/10.1186/s12875-015-0348-0>
27. Walsh A, Edwards H, Fraser J. Influences on parents' fever management: beliefs, experiences and information sources. *J Clin Nurs* 2007;**16**:2331–40. <https://doi.org/10.1111/j.1365-2702.2006.01890.x>
28. Kai J. What worries parents when their preschool children are acutely ill, and why: a qualitative study. *BMJ* 1996;**313**:983–6. <https://doi.org/10.1136/bmj.313.7063.983>
29. Geijer H, Udumyan R, Lohse G, Nilsagård Y. Temperature measurements with a temporal scanner: systematic review and meta-analysis. *BMJ Open* 2016;**6**:e009509. <https://doi.org/10.1136/bmjopen-2015-009509>
30. Van den Bruel A, Haj-Hassan T, Thompson M, Buntinx F, Mant D, European Research Network on Recognising Serious Infection investigators. Diagnostic value of clinical features at presentation to identify serious infection in children in developed countries: a systematic review. *Lancet* 2010;**375**:834–45. [https://doi.org/10.1016/S0140-6736\(09\)62000-6](https://doi.org/10.1016/S0140-6736(09)62000-6)
31. Sollai S, Dani C, Berti E, Fancelli C, Galli L, de Martino M, Chiappini E. Performance of a non-contact infrared thermometer in healthy newborns. *BMJ Open* 2016;**6**:e008695. <https://doi.org/10.1136/bmjopen-2015-008695>

# Appendix 1 Participant characteristics

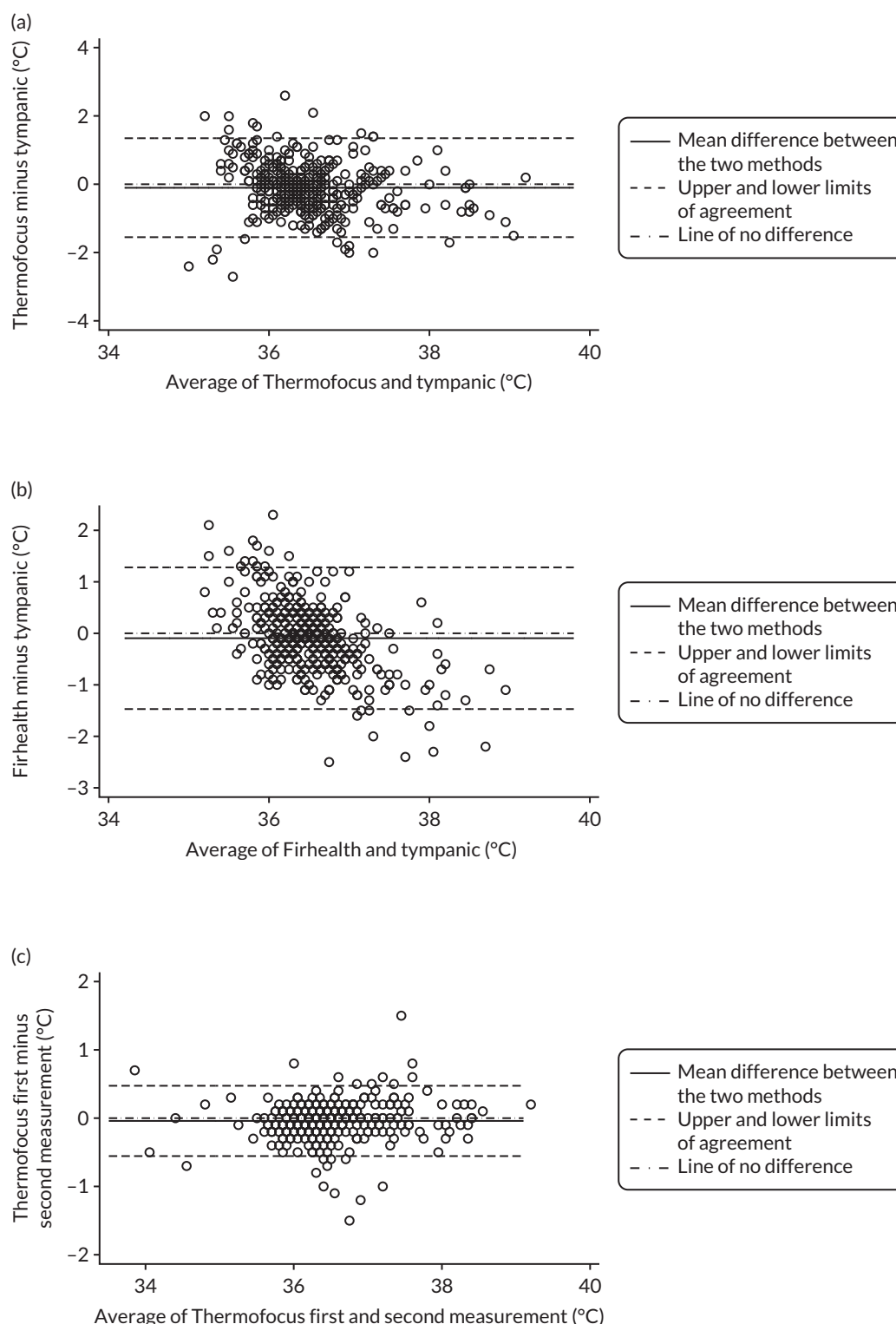
TABLE 6 Qualitative interview participant characteristics

Characteristics	Number
<b>Parental characteristics</b>	
Gender	
Male	6
Female	15
Age (years)	
20–30	3
31–40	11
> 40	3
White British ethnicity	
Yes	15
No	6
<b>Case characteristics</b>	
Age of child	Range 4 months–5 years
History of fever this episode	
Yes	6
No	15
Number of children in household	
1	12
2	9

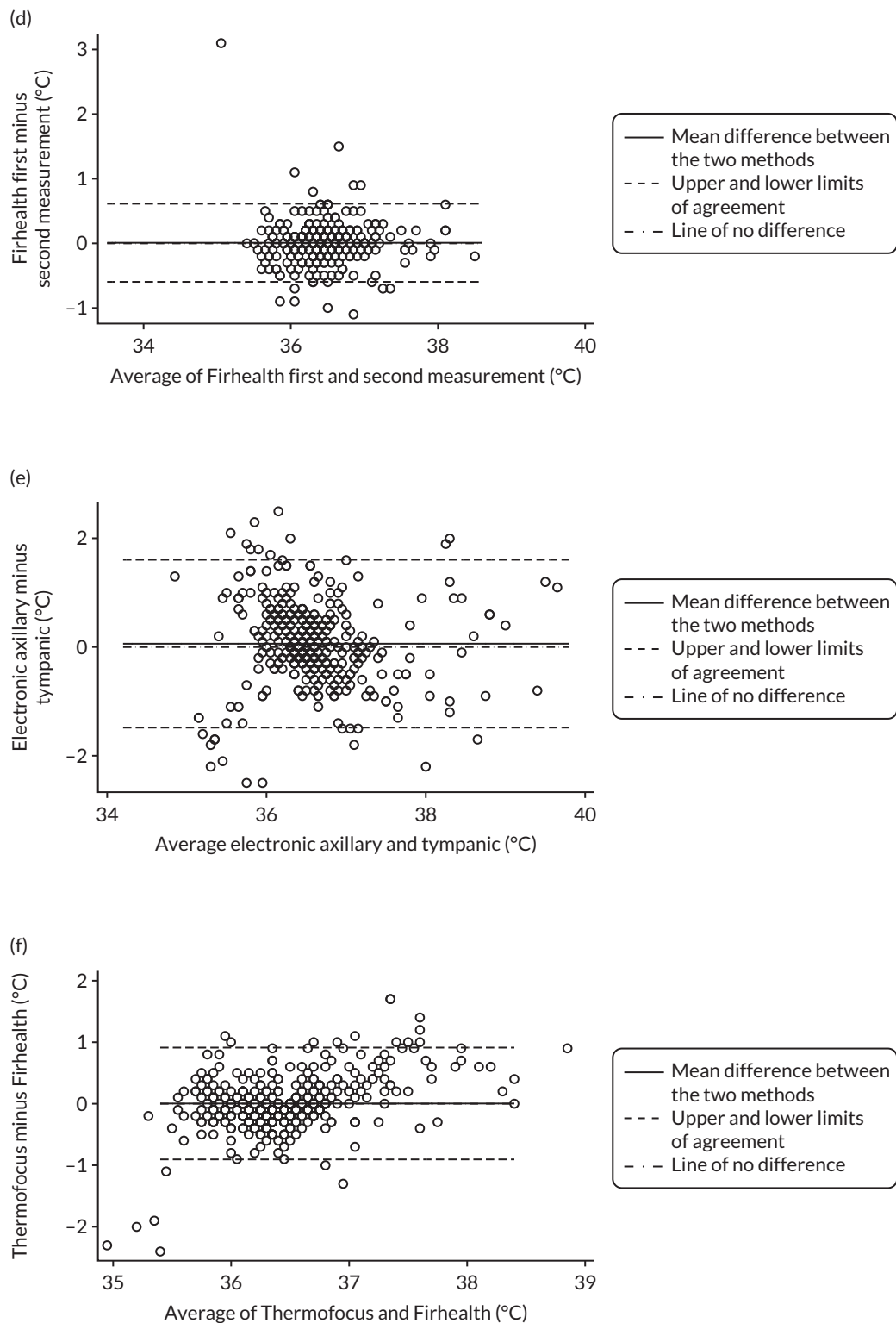




## Appendix 2 Bland–Altman plots



**FIGURE 4** Bland–Altman plots of agreement for the other comparisons. Agreement between (a) Thermofocus and tympanic; (b) Firhealth and tympanic; (c) Thermofocus first and second measurements; (d) Firhealth first and second measurements; (e) electronic axillary and tympanic; and (f) Thermofocus and Firhealth. (*continued*)



**FIGURE 4** Bland-Altman plots of agreement for the other comparisons. Agreement between (a) Thermofocus and tympanic; (b) Firhealth and tympanic; (c) Thermofocus first and second measurements; (d) Firhealth first and second measurements; (e) electronic axillary and tympanic; and (f) Thermofocus and Firhealth.



EME  
HS&DR  
HTA  
PGfAR  
PHR

Part of the NIHR Journals Library  
[www.journalslibrary.nihr.ac.uk](http://www.journalslibrary.nihr.ac.uk)

*This report presents independent research funded by the National Institute for Health Research (NIHR).  
The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the  
Department of Health and Social Care*

***Published by the NIHR Journals Library***