Clinical Assessment of Non-Contact Infrared Thermometers as a medical countermeasure

Stacey J.L. Sullivan[1], Jean E. Rinaldi[1], Prasanna Hariharan[1]*, Jon P. Casamento[1], Seungchul Baek[2], Nathanael Seay[2], Oleg Vesnovsky[1], L.D. Timmie Topoleski[2]

[1]Division of Applied Mechanics, Office of Science and Engineering Laboratories, Center for Devices and Radiological Health, United States Food and Drug Administration,
[2]University of Maryland Baltimore County

Abstract

Background: Non-contact infrared thermometers (NCITs) are widely used during disease outbreaks as a temperature-measurement tool for screening the general public, travelers at ports of entry, and isolating patients in healthcare and other settings.

Methods: A clinical study was conducted with 1113 adult subjects using six different commercially available NCIT models to assess their temperature measurement accuracy. A total of 60 NCITs were tested with 10 units for each model. The NCIT-measured temperature was compared with the oral reference temperature.

Results: The mean difference between the oral reference thermometer and NCIT measurement (clinical bias) was different for each NCIT model. The clinical bias ranged from just under -0.9 °C (under-reporting) to just over +0.2 °C (over-reporting). The majority of the individual measurement differences ranged between -2 °C and +1 °C with extreme cases ranging from -3 °C to +2 °C. Depending upon the NCIT model, 48% to 88% of the individual temperature measurements were outside the labeled accuracy stated by the manufacturers. The sensitivity, which ranges from 0 (no detection) to 1 (ideal detection), of the NCIT models for detecting a subject's temperature above 38 °C ranged from 0 to 0.69.

Conclusions: Overall, our results indicated that the tested NCIT devices may not be consistently accurate enough to determine if a subject's temperature exceeds a specific threshold (e.g., 38 °C). Inter-model variability and intra-model accuracy of the displayed temperature were found to be outside of acceptable limits. Accuracy and credibility of NCITs should be more thoroughly investigated in future studies before considering them as an effective screening tool.

Introduction

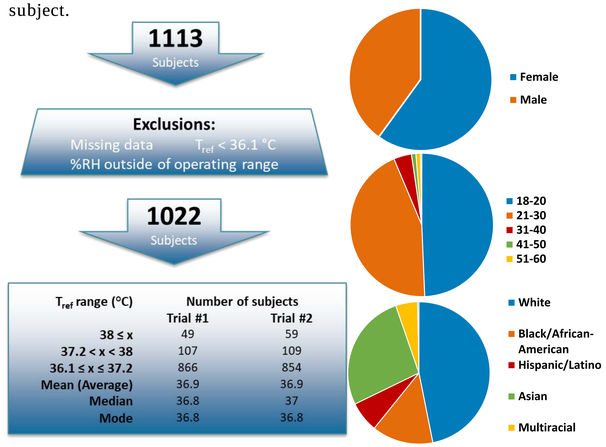
Non-contact Infrared Thermometers (NCITs) are being used as a temperature measurement tool for screening and isolating potentially infected people with elevated temperature in healthcare settings, ports of entry and in other group settings. NCITs do not measure the core body temperature directly but are designed to correlate with a reference body site temperature, such as the oral temperature. The forehead skin surface temperature is measured based upon detection of infrared radiant energy from the surface of the skin. The temperature of the forehead skin surface is lower than reference body site temperature. Therefore, manufacturers typically use a proprietary algorithm and hardware design features to compensate for the difference between the forehead skin surface temperature and the reference body site temperature - the "adjusted mode," typically referred to as "subject mode" for most NCITs. The algorithm used to adjust the temperature also may compensate for other factors such as variations in room temperature, skin emissivity, and clinical and hardware biases.

Although NCITs may be the primary tool for temperature screening, clinical studies have reported mixed performance in terms of their accuracies. The accuracy of NCITs are currently evaluated using the ASTM E1965 and ISO 80601-2-56 standards. Both standards set a requirement for the laboratory error to be within ±0.3 °C. The objective of this study was to evaluate, analyze, and report the accuracy of multiple units of various commonly-available NCIT models in a large-scale controlled clinical study comprised of both afebrile and febrile adult subjects.

Materials and Methods

Six different commercially available NCIT models from different manufacturers that measure temperature at the center of the forehead were tested. Ten units of each model were purchased from commercial vendors. NCITs were divided into ten identical sets; each measurement set contained one unit of each of the six different NCIT models, labeled A through F. Thermometers were cleaned and prepared according to the manufacturer's instructions for use and had fresh batteries installed prior to testing.
Temperature measurement sequence:

All measurements were taken at the University of Maryland Health Center. Room temperature and humidity were recorded at the time of the measurements. The same operator made the all measurements for a single subject.



T_{ref} range (°C)	Number of subjects Trial #1	Trial #2
38 ≤ x	49	59
37.2 < x < 38	107	109
36.1 ≤ x ≤ 37.2	866	854
Mean (Average)	36.9	36.9
Median	36.8	37
Mode	36.8	36.8

Results

Table 1. NCIT model performance evaluated to specification.

NCIT Models	A		B		C		D		E		F	
Stated accuracy for measurement range	±0.2°C (36°C to 39°C) ±0.3°C (<36°C;>39°C)		±0.2°C		±0.2°C (36°C to 39°C) ±0.3°C (<36°C;>39°C)		±0.2°C		±0.3°C		±0.2°C (36°C to 39°C) ±0.3°C (<36°C;>39°C)	
Trial	#1	#2	#1	#2	#1	#2	#1	#2	#1	#2	#1	#2
Total # of readings	1021	1022	1022	1022	1022	1022	884	884	1019	1019	886	886
# outside accuracy	493	523	503	497	606	538	564	527	874	891	557	545
% outside accuracy	48.3	51.2	49.2	48.6	59.3	52.6	63.8	59.6	85.8	87.4	62.9	61.5

Table 2. NCIT model clinical bias (mean error) and distribution.

$\Delta T = T_{NCIT} - T_{Ref}$ (°C)												
NCIT Models	A		B		C		D		E		F	
Trial	#1	#3	#1	#3	#1	#3	#1	#3	#1	#3	#1	#3
Clinical bias (Ave)	-0.23	-0.28	-0.22	-0.24	0.15	0.14	-0.32	-0.31	-0.87	-0.89	0.21	0.23
Standard Deviation	0.46	0.45	0.43	0.40	0.42	0.39	0.61	0.54	0.53	0.54	0.48	0.43
95th percentile	0.40	0.30	0.40	0.30	0.70	0.70	0.50	0.40	-0.10	-0.20	0.90	0.80
5th percentile	-1.10	-1.20	-1.00	-1.00	-0.70	-0.60	-1.50	-1.30	-1.80	-1.90	-0.60	-0.50

Results

For all models, more than 48% of the clinical measurements fell outside of the manufacturer's labeled accuracy (Table 1). The clinical bias per ASTM E1965 (Trial #1) ranged from under-reporting the temperature by -0.87 °C to over-reporting the temperature by 0.21 °C. Overall, model E had the greatest clinical bias (-0.89 °C) while Model C had the least clinical bias (0.14 °C). All six NCIT models had relatively large standard deviations compared to mean (Table 2 and Figure 1). The 5th percentile value for ΔT was between -1.9 °C and -0.5 °C. For the six NCIT models, the mode value for ΔT varied between -0.7 °C and 0.4 °C (Figure 2). The correlation between ΔT and T_{ref} showed that the difference, ΔT, changed as a function of T_{ref} for all NCIT models (Figure 3) for Trial #1.

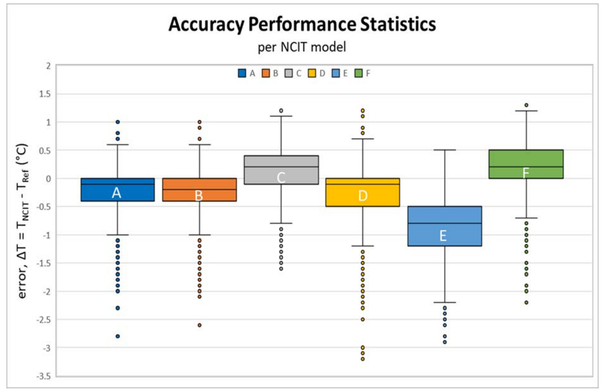


Figure 1. Accuracy performance statistics for each NCIT model for Trial #1. The midline indicates the median, the box top captures 25% of the data above the median and the box bottom captures 25% of the data below the median. The whiskers (error bars) represent that maximum and the minimum ΔT. The circles represent outlier data.

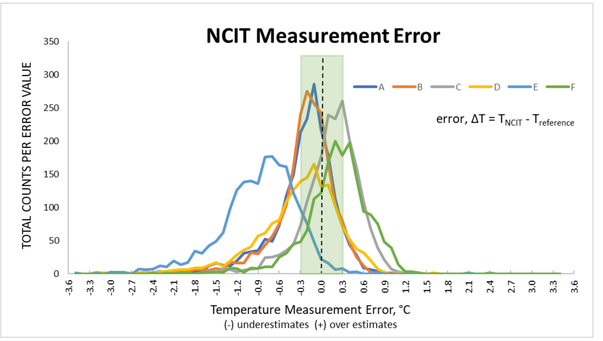


Figure 2. Total counts per error value, per NCIT model for both Trial #1 and #2. Green area indicates ±0.3 °C laboratory accuracy zone per the standards; dashed black line indicates the zero error line.

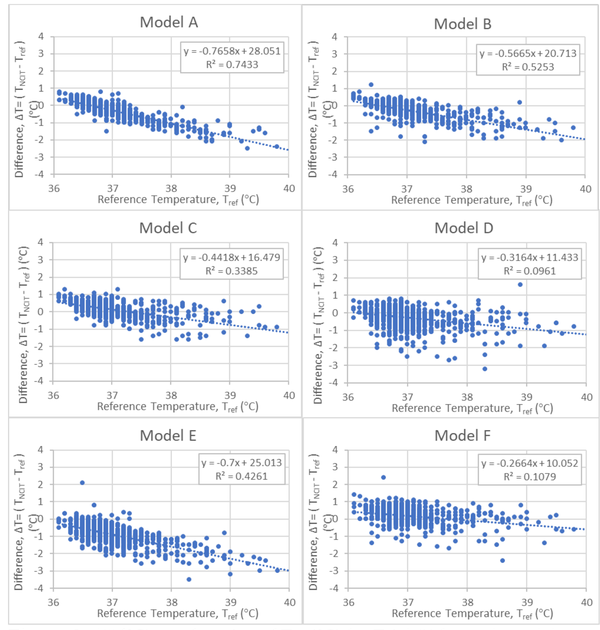


Figure 3. ΔT (°C) as a function of T_{ref} (°C) for Trial #1 for each model.

Discussion & Conclusion

Model-to-model variability and individual model accuracy in the displayed temperature were found to be outside of acceptable limits.

Our study protocol was designed to minimize the inaccuracies due to user error and environmental factors. In a real-world setting (e.g., transit centers, pre-clinical triage, and other screening locations), the additional inaccuracies and variabilities will only increase the error in NCIT-measured body temperature. Therefore, it is critical to follow the manufacturer's instructions for use to minimize inaccuracies due to user error and environmental factors and to optimize and ensure proper device performance.

Overall, our results indicate that some NCIT devices may not be consistently accurate enough to be used as a stand-alone temperature measurement tool to determine if the subject's temperature exceeds a specific threshold (e.g., 38 °C) in an adult population.

Acknowledgements & Disclosures

This clinical study was conducted under an FDA Institutional Review Board (IRB) approved protocol.

The findings and conclusions in this study have not been formally disseminated by the Food and Drug Administration and should not be construed to represent any Agency determination or policy.

Funding for this study was provided by FDA's Medical Countermeasures Initiative.