Comparison of ARC InstaTemp MD® versus the Welch Allyn SureTemp 692 & Covidien Genius II measuring Repeatability, Clinical Bias, and Line of Agreement across four patient cohorts.

Steven Gerst, MD, MBA, MPH, CHE, Finbar Dolan PhD, CEng, Martin Crawley, B.Eng. Mark Khachaturian, PhD

Abstract:

This is a cross sectional clinical observational study designed to compare the accuracy and reliability of the ARC Devices, Ltd. ARC InstaTemp MD® Infrared, Non-Touch thermometer versus a current industry standard (a) Welch Allyn SureTemp® Plus 692 *sublingual probe thermometer and (b) Covidien* Genius™ 2 *tympanic membrane thermometer. The study was performed to meet ISO standard 80601-2:56:2009 section 201.102 to validate clinical accuracy on 136 patients of which 38.4% were febrile. Subjects were divided into 4 age groups: < 3 mos. old, 3 mos. -1 yr. old, 1 yr.-5 yrs. and >5yrs. old.*

The results demonstrate that the ARC Devices, Ltd. ARC InstaTemp MD® Infrared, Non-Touch thermometer accurately and consistently estimates core body temperature in a large representative hospital population. The InstaTemp MD® has significantly better precision than both the Welch Allyn and Covidien tympanic membrane thermometer that are in widespread use in hospital practice both in the US and EU.

Background:

The ARC Devices, Ltd. ARC InstaTempMD ® digital infrared, Non-Touch thermometer estimates Core Body temperature of heat emissions from the patient's skin surface in the center of the forehead above the eyebrows. It affords reliable, accurate and safe hyperthermic fever detection as well as hypothermic determinations without direct body contact. This was an independent research study conducted by the Clinical Research Team, HRB Clinical Research Facility - Cork, affiliated with Mercy University Hospital, Cork, Ireland.

Methodology:

The Independent Clinical Research Team ("CRF-C") of HRB Clinical Research Facility – Cork conducted this cross-sectional, clinical investigation to quantify the reliability and validity of body temperature estimated from emitted surface body heat as measured by the ARC Devices, Ltd. *ARC InstaTemp MD®* Infrared, Non-Touch thermometer (referred to herein as *ARC InstaTemp MD®*). The study was funded by ARC Devices; the independent research team was led by Prof Joe Eustace, Director of the HRB CRF-C in collaboration with Dr. Conor Deasy, Dr. Louise Gibson and Dr. Pat Barry, consultant Physicians at Cork University Hospital and Dr. Darren Dahily Senior

Comparison of ARC InstaTemp MD® versus the Welch Allyn SureTemp 692 & Covidien Genius II measuring Repeatability, Clinical Bias, and Line of Agreement across four patient cohorts.

Steven Gerst, MD, MBA, MPH, CHE, Finbar Dolan PhD, CEng, Martin Crawley, B.Eng. Mark Khachaturian, PhD

Lecturer in Biostatistics at the CRF-C.. The research team maintained full editorial power over the study report which was submitted to the FDA.

Triplicate temperature readings using the *ARC InstaTemp MD*® were compared in 136 patients to concurrent measurements obtained using Reference Clinical Thermometry units commonly used by hospitals, clinics, ambulatory surgical centers, physicians' offices and other clinical locations in both the United States and in Europe (i.e., the sublingual Welch Allyn oral probe thermometer and/or the Covidien tympanic membrane thermometer).

Patients were recruited during the late spring and early summer from general medical wards, pediatric wards, the Acute Medical Unit and the Emergency Department of Cork University Hospital, a large tertiary referral in University Hospital in Cork, Ireland. The Study was conducted by The HRB Clinical Research Facility – Cork, an independent Research Facility.

The study was conducted under the HRB Facility's Standard Operating Procedures and Quality Program. The Clinical Research Team ("CRF-C") study team take full responsibility for the accuracy of the study data, and the resulting analysis of the results in this whitepaper. Full editorial power over the final analysis and report remained with the study team.

Study Procedures:

The study protocol was approved by the Clinical Research Ethics Committee of the Cork Teaching Hospitals. Informed written consent was obtained from all adult subjects or, in the case of children, from the parent or legal guardian, along with the age appropriate assent of the child. Consenting subjects who were admitted to Cork University Hospital were eligible to be enrolled in the study.

Subjects who were being treated with (a) cooling blankets, (b) fans, (c) corticosteroids, (d) thyroxine replacement therapy, (e) barbiturates, (f) had eaten or ingested fluids in the previous 15 minutes, (h) were unable to expose their forehead, (i) had bilateral in-situ hearing aids, or (j) had active infection or inflammation at the proposed measurement site were excluded. Having obtained informed consent, a unique study identifier was applied to the patient's data. This tracking number was used exclusively as the identifier on study case report forms.

Patient temperatures were measured as per the manufacturer's instructions using (1) the *ARC InstaTemp MD*®, (2) the CE marked Covidien *Genius™ 2* Tympanic Thermometer (in routinely used in hospital practice in Ireland), and (3) the CE marked Welch Allyn *SureTemp® Plus 692* sublingual thermometer which is widely used in the United States.

Comparison of ARC InstaTemp MD® versus the Welch Allyn SureTemp 692 & Covidien Genius II measuring Repeatability, Clinical Bias, and Line of Agreement across four patient cohorts.

Steven Gerst, MD, MBA, MPH, CHE, Finbar Dolan PhD, CEng, Martin Crawley, B.Eng. Mark Khachaturian, PhD

The *ARC InstaTemp MD*® was provided direct from ARC Devices from a validated product manufacturing line. The Covidien *Genius™ 2* and the Welch Allyn *SureTemp® Plus REF 692* standard thermometers were purchased commercially immediately prior to, and specifically for, the purpose of the study. For the two standard probe thermometers, a new probe cover was used for each measurement attempt.

For children under 5 years of age, readings were only taken from a single Reference Clinical Thermometer using the Covidien Genius™ 2 tympanic method as is the routine clinical practice at Cork University Hospital. Up to three sets of readings were taken with each method, as tolerated by each patient. Measurements were stopped for any patient who became distressed or who did not wish to continue for any reason. A parent or legal guardian was present throughout the procedure for all infants and children.

Avoiding Bias:

To avoid introducing any study bias into the research protocol, readings were obtained using two alternating thermometer order sequences. Sequence A were temperatures taken from patients in the sequential order: (1) Sublingually (using the Welch Allyn *SureTemp® Plus 692*), then (2) using *ARC InstaTemp MD*® digital, "Non-Touch" thermometer, and finally (3) using the Covidien *Genius™ 2* Tympanic Membrane thermometer.

Sequence B was: (1) start by taking the patient's temperature using the Covidien Genius™ 2 Tympanic Membrane thermometer, then, (2) using the *ARC InstaTemp MD*® digital, "Non-Touch" thermometer, and finally (3) sublingually (using the Welch Allyn *SureTemp® Plus 692*).

Whenever possible, tympanic readings were taken in alternating ears and the ear sequence was recorded. Two minutes were allowed to lapse prior to a second reading in the same ear. Measurement time was recorded in HH:MM:SS (hours, minutes, and seconds). Ambient temperature was also recorded using a temperature meter.

All temperatures were recorded in Fahrenheit. A subject was considered febrile if at least 2 readings using the method in routine hospital use, (i.e. tympanic) were ≥100.4°F. Data was separately double entered into an analysis database, with any discrepancies reconciled against the original case report form. Analyses were performed using R statistical programming language, version 3.2.5 (2016-04-14), running on 64-bit Windows.

The recruitment of febrile infants < 3 months into the trial was particularly challenging as critically ill infants were treated in the Neonatal Intensive Care Unit in temperature controlled enclosed cots,

while the majority of infants admitted on the general pediatrics wards who had a history of fever had already commenced (and were maintained on antipyretics) and in most cases remained afebrile by study criteria post admission.

Clinical Accuracy Validation ISO 80601-2-56:2009. Section 201.102

The disclosure of clinical accuracy validation of a thermometer under test (TUT) against a reference clinical thermometer (RCT) should include its *clinical bias* (Δ_{cb}), its *limits of agreement* (L_A), and its *clinical repeatability* (σ_r).

This validation was carried out in each of the following applicable age groups:

Age Group Cohort	Subgroups
Group A - < 1 year old	**Subgroup A1**: 0 <3 months **Subgroup A2**: 3 months to 1 yr.
Group B – 1yr. to 5 yrs. old	
Group C - > 5 years old.	

As per the ISO 80601-2-56:2009. Section 201.102, Clinical Accuracy validation, any group should include at least 35 subjects, with at least 15 subjects in any subgroup. Each age group should contain between 30% and 50% febrile subjects. The total sample across all groups should be a minimum of 105 subjects.

Patient Cohort Characteristics of this Study

To meet the guidelines:

Age Group Cohort	Number of Patients Studied	Percent Febrile
< 3 mos. old	19	5% (n=1)
3 mos. – 1 yr. old	22	45% (n=10)

Total < 1 yr. old.	**39**	**26.8% (n=11)**
1 yr. – 5 yrs. old	46	41% (n=19)
>5 yrs. old	51	45% (n=23)
Total Patients	**136**	**38.4% (n= 53)**

Clinical Bias

Clinical bias is the difference in the first TUT measurement and the first RCT measurement, averaged across all subjects for each age group.

$$\Delta_{cb} = \frac{\sum_{i=1}^{n} TUT_i^t - RCT_i^t}{n}$$

where *i* is the index number for each subject, *n* is the total number of subjects in the age group, and TUT_i^t, RCT_i^t are the first output measurements taken with the TUT and RCT.

Welch Allyn measurements were only taken in the 5+y age group. The clinical bias for the *ARC InstaTemp MD*® compared with the Welch Allyn in this age group was 0.07 °F.

Covidien measurements were collected for all four age groups. The clinical bias for the *ARC InstaTemp MD*® compared with the Covidien in each age group was as follows:

Age Group	**Clinical Bias *ARC InstaTemp MD® vs the Covidien Genius 2 Tympanic Membrane Thermometer***
< 3 months old	0.14 °F
3 months to1 yr. old	0.05 °F
1 yr. to 5 yrs. old.	-0.72 °F

> 5 yrs. old.	-0.55 °F

Limits of Agreement

The Limits of Agreement are calculated as two times the standard deviation of Δ_{cb} for each age group.

$$L_A = 2 \times \sigma_{\Delta_{cb}}$$

, where $\sigma_{\Delta_{cb}}$ is calculated as:

$$\sigma_{\Delta_{cb}} = \sqrt{\frac{\sum_{i=1}^{n}[\,(TUT_i^t - RCT_i^t) - \Delta_{cb}]^2}{n-1}}$$

Welch Allyn measurements were only taken in the 5+yr. age group. The limits of agreement for the *ARC InstaTemp MD*® compared with the Welch Allyn in this age group was 1.96 °F.

The Bland-Altman plot for the *ARC InstaTemp MD*® vs. the Welch Allyn thermometer is given overleaf, **Figure 1**.

Comparison of ARC InstaTemp MD® versus the Welch Allyn SureTemp 692 & Covidien Genius II measuring Repeatability, Clinical Bias, and Line of Agreement across four patient cohorts.

Steven Gerst, MD, MBA, MPH, CHE, Finbar Dolan PhD, CEng, Martin Crawley, B.Eng. Mark Khachaturian, PhD

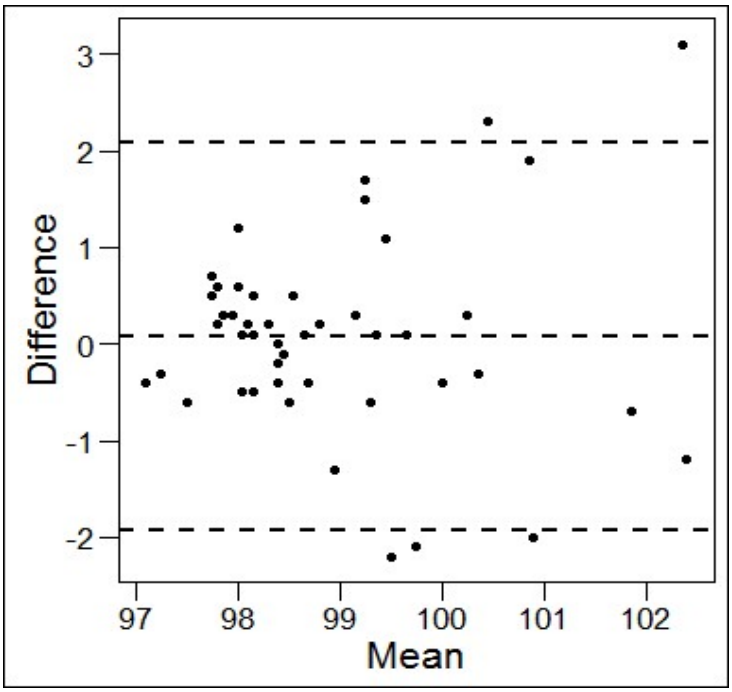

Fig 1. Bland-Altman plot for the *ARC InstaTemp MD*® vs. the Welch Allyn thermometer.

Covidien measurements were collected for all four age groups. The limits of agreement for the *ARC InstaTemp MD*® compared with the Covidien in each age group was as follows:

Age Group	**Limits of Agreement *ARC InstaTemp MD® vs the Covidien Genius 2 Tympanic Membrane Thermometer***
< 3 months old	2.15 °F
3 months -1 yr. old	2.79 °F
1 yr. – 5 yrs. old.	2.41°F
> 5 yrs. old.	2.88 °F

Comparison of ARC InstaTemp MD® versus the Welch Allyn SureTemp 692 & Covidien Genius II measuring Repeatability, Clinical Bias, and Line of Agreement across four patient cohorts.

Steven Gerst, MD, MBA, MPH, CHE, Finbar Dolan PhD, CEng, Martin Crawley, B.Eng. Mark Khachaturian, PhD

Clinical Repeatability

Clinical repeatability is calculated for all subjects, across all age groups, excluding febrile subjects under the age of 5 years old. It is calculated as the pooled standard deviation of triplicate measurements of the TUT. The standard deviation for each subject, σ_j, is calculated as follows:

$$\sigma_j = \sqrt{\frac{\sum_{i=1}^{m}(TUT_i^t - \overline{TUT_j^t})}{m-1}}$$

where $\overline{TUT_j^t}$ is the average of all measurements for subject *j*, and *m* is the number of measurements for the subject. The standard deviations are then pooled across all subjects as follows:

$$\sigma_r = \sqrt{\frac{\sum_{j=1}^{N}\sigma_j^2}{N}}$$

where *N* is the total number of subjects across all age groups.

The clinical repeatability of the *ARC InstaTemp MD*® in our subject sample was 0.21, omitting the febrile under-5-year-olds (n = 28).

The repeatability values in the table below were derived based on calculating the average of the standard deviation of the three repeat measurements of the respective thermometers for the entire population of patients in the study (*ARC InstaTemp MD*®, Welch Allyn SureTemp 692, Covidien Genius II)

Device	Repeatability (3 measurements on same patient)
Instatemp MD	± 0.18 [ºF]
Welch Allyn	± 0.35 [ºF]
Covidien	± 0.38 [ºF]

Comparison of ARC InstaTemp MD® versus the Welch Allyn SureTemp 692 & Covidien Genius II measuring Repeatability, Clinical Bias, and Line of Agreement across four patient cohorts.

Steven Gerst, MD, MBA, MPH, CHE, Finbar Dolan PhD, CEng, Martin Crawley, B.Eng. Mark Khachaturian, PhD

Additional Analyses

Covidien vs. Welch Allyn

For comparison, we also evaluated the Covidien (as the TUT) vs. the Welch Allyn (RCT) in the 5+y age group. The clinical bias for the Covidien was 0.59 °F (vs. 0.07 °F for the *ARC InstaTemp MD*®), and the limits of agreement was 2.37 °F (vs. 1.96 °F. for the *ARC InstaTemp MD*®).

Clinical Repeatability for other thermometers

In comparison, the clinical repeatability in subjects older than 5 years using the Welch Allyn oral thermometer was 0.47. The clinical repeatability in the full sample was 0.44 for the Covidien (vs. 0.21 for the *ARC InstaTemp MD*®).

Conclusion

The above study in a large representative hospital population reveals *the ARC InstaTemp MD*® Infrared, Non-Touch thermometer to have equal levels of reproducibility and accuracy which align with existing FDA-cleared thermometers that are in widespread hospital practice.