Non-invasive and wearable early fever detection system for young children

MDP Garcia-Souto[1,2] and P Dabnichki[1,3]

[1]School of Engineering and Materials Science, Queen Mary, University of London
Mile End road, London E1 4NS, UK
[2]Present address: Medical Physics and Biomedical Engineering, University College London
Malet Place Engineering Building - Gower Street, London WC1E 6BT, UK
[3]Present address: School of Engineering, RMIT University
124 La Trobe St, Melbourne VIC 3000, Australia

Corresponding address: p.garciasouto@ucl.ac.uk

Abstract

Fever in young children is taken seriously by healthcare professionals as it indicates an underlying infection which can be life-threatening. Core body temperature can be accurately measured using traditional techniques, but these are not suitable for non-invasive monitoring during normal life. This study investigates the possibility of fever monitoring in children under 2 years of age in a non-clinical setting based on various local skin temperatures. Various system designs are presented, i.e. single vs multi-sensor systems, and a set of sensors either localized or distributed across the body. The probability of positive fever identification on feverish children ranges from ~40% to 77% using 1 and 5 sensors respectively, while the detected false positives are a 10%. We conclude that a continuous and non-invasive fever monitoring in children under 2 years is possible by the propose method, providing a suitable solution for early fever detection and alert.

Keywords
Child; fever; core temperature; skin temperature; sensor.

1 Introduction

Fever in young children usually indicates an underlying infection, caused either by a self-limiting virus or a more dangerous bacterium, such as meningitis or pneumonia which can be life-threatening as pointed-out by the National Collaborating Centre for Women's and Children's Health [1]. In fact, infections remain the leading cause of death in children under the age of 5 years. Fever is very common in young children, with between 20 and 40% of parents reporting it each year, probably the commonest reason for a child to be taken to the doctor, and the second most common reason for a child being admitted to hospital. Even when the fever is reported by the parents or carers, fever ought to be taken seriously by healthcare professionals. The *NICE Clinical Guideline 160* (CG160) applies to children under 5 years old [1]. CG160 states that healthcare professionals are required to set a face-to-face visit when the children have any features belonging to the amber / intermediate risk category (e.g. pallor on skin, lips o tongue; no smile or decrease activity; nasal flaring; poor feeding in infants; etc.) though the urgency can be determined remotely. However, in the case of children presenting features within the red / high risk category (e.g. pale colour of skin, lips or tongue; weak, high-pitched or continuous cry; grunting; etc.), the healthcare professional needs to see the children within 2 hours even if they are not considered to have an immediately life-threatening illness [1].

Core body temperature can be accurately measured by using a range of traditional 'gold standard' techniques such as pulmonary artery temperature, oesophageal temperature, intra-abdominal temperature and rectal temperature in clinical settings, or oral temperature, axilla temperature or tympanic temperature when a practical and less-invasive option is required and sufficient. In particular, in non-clinical settings the ear temperature measured with an infrared device is the most popular alternative to the 'gold standards', as it is easy to apply, fast, safe and highly tolerable by the patient [2,3], while it has been found accurate both during febrile and non-febrile periods [4], and valid to assesses the presence of fever [5], and it is one of the recommended methods to be used with children between 4 weeks and 5 years of age. However, all these methods are typically used either in clinical settings or at home when other signs suggest the presence of fever, and are not suitable for long-term monitoring in normal-life situations. This is not sufficient to monitor babies and children for fever occurrence, especially in schools or nurseries where the ratio of children to adult is considerable as children depend on their carer to identify such condition. Therefore, the study and development of a non-intrusive indicator that provides reliable means for fever detection in children is of practical importance.

The advances of technology, sensors, smartphones and smart watches, and communication systems such as connectivity with The Internet of Things have supported the development of monitoring systems to improve health-care in our daily-life. These systems are being developed for the monitoring of a variety of physiological signals to inform doctors or to raise an alarm in case of abnormalities or danger, e.g. body posture and falls monitoring [6,7],

cardiovascular related signals [6,8], respiratory problems [6], thermal comfort [9], temporal pattern of perspiration [10], skin temperature [11], or mental stress via skin conductance [12].

Literature also shows that there is a need for non-invasive core temperature and fever monitoring systems that can be used in non-clinical environments. These current systems are typically based on temperature measurements of the skin, and can be used in a variety of circumstances, e.g. providing core temperature measurements in real-time during remote consultations with doctors [8], detecting moisture and core temperature of bedridden patients via sensors embedded in their garments [13], fever detection at the entrance of a class during infectious outbreaks [14], or a wrist-band for all purposes fever alarm [15]. The accuracy is obviously limited as skin temperature depends on many factors [16,17] and only provides an approximation of the core temperature. Despite all those shortcomings, some authors have successfully used skin temperature at the forehead to determine core temperature within ±0.5°C of distal oesophageal temperature for intensive care patients [18], demonstrating the potential of the method to substitute the more accurate invasive procedures.

However, typically sensors that measure and monitor core temperature via skin temperature involve electronics and may require considerable power supply so while alternatives to AC power are currently sought [19], and have benefits especially for infants. Furthermore, compounded data for children are limited, with just a few studies providing significant data related to the skin and core temperature of children [20] and babies below two years of age while the characteristic skin temperature mapping is different than that of adults [17].

The aim of this study is to investigate if skin temperature can be used to monitor fever in a continuous fashion on children under 2 years of age in a non-clinical setting, and if affirmative, describe the suitable system. As current fever and thermal comfort monitoring devices are developed for adults, the work directly addresses the existing gap in the acquired data and methodologies used to develop a system relevant for infants. Early detection of fever could potentially save youngster's life to conditions such as meningitis, pneumonia and other life threatening conditions.

2 Study design

## 2.1 Sensor requirements and system design

A novel non-clinical fever monitoring method for children under 2 years of age is sought based on skin temperatures at different locations on the body surface. The criteria measurement strategy for this specific type of population needs to be proposed based on extensive experimental studies. The system should offer a suitable non-invasive solution for early detection and alert.

For practical and safety reasons, the designed system/sensor should be embedded in the clothing (**Figure 1**), and should not use wires or batteries. A binary output would be

sufficient, indicating presence of fever or no-fever (**Figure 2**). A thermochromic pigment applied over a selected area of the fabric could be used. Identification of the suitable skin locations and skin temperature threshold value matching with the onset of fever is investigated and reported in this study.
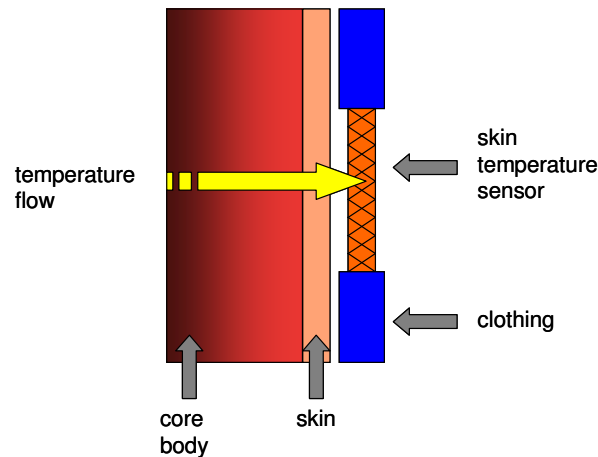
Figure 1: Skin temperature sensor applied over the skin
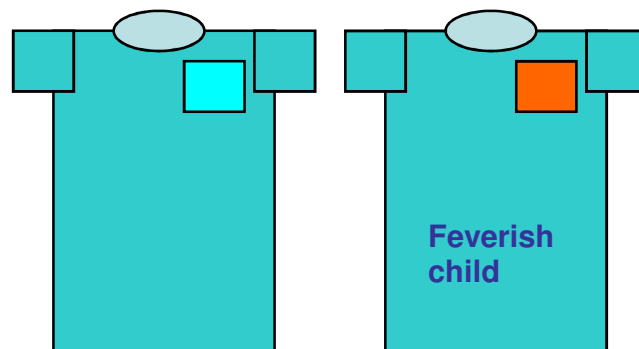
Figure 2: Possible location and outputs of the fever sensor on the garment

Various methods to ascertain core temperature were considered. However, working with this particularly vulnerable subject group strongly limits the measurement techniques that could be used, with typical 'gold standards' being not allowed. On the one hand, the "*NICE Clinical Guideline 160- Fever in under 5s: assessment and initial management –*" recommends the use of an electronic thermometer in the axilla, a chemical dot thermometer in the axilla*,* or an IR ear thermometer to obtain an approximate core temperature value in children aged 4 weeks to 5 years [1]. The first two alternatives (axilla thermometers) were deemed less practical and of lower performance than the IR ear thermometer [21]. On the other hand, there are concerns in the literature with the use of IR ear thermometers in clinical settings or as a substitution of a 'gold standard' [22,23], while others believe it correlates well with 'gold-standards' such as pulmonary artery temperature [21] and rectal temperature [2,4,5,24-26]. However, this disagreement mainly focuses on the use of tympanic temperature as a substitute of a 'gold standards' or in clinical and intensive care situations. Instead, the

suitability of IR ear thermometers in non-clinical settings is much more popular and recommended, particularly in studies with children [4,5,24,26]. Finally, it is necessary to ascertain the core temperature fairly for the design of this system, but precision was not critical as it is in clinical settings. This is because our system only aims to work as a fever alert and not as a fever diagnostic method; and also the largest inherent error from our system comes from the skin temperature variability. Considering all these aspects, it was decided that an IR ear thermometer will be used as to get an approximation of the core temperature, and that it is a sufficient and valid option.

2.2 Data collection

The children cohort was recruited and tested following a study protocol approved by the University of London Ethical Committee (QMREC2008/72). A brief description of the experimental protocol is given here, but full details were reported by Garcia-Souto and Dabnichki previously [17] along with the group stratification.

A total of 138 subjects participated in the study. They were gender-balanced, 3 to 24 months of age, and with BMI between 14.4 and 21.8 kg/m^2. Core temperature and 16 local skin temperatures (the 14 points recommended in ISO 9886 [27] plus lateral neck and left wrist) (**Figure 3**) were measured in 540 occasions, deeming a total of 9180 samples. Volunteers' gender, age, Body Mass Index (BMI), clothing and their status (awake or asleep) were recorded.

All measurements were taken indoors at the nurseries after a minimum of 15 minutes allowing to reach stable skin temperature levels [28]. Room temperature and humidity was maintained within a narrow range (23.7°C (SD=1.7°C) and 55.1% (SD=6.8%)) with no significant air velocity or radiant temperature from either direct exposure to the sun and/or heating elements.

Tympanic temperature, used as core temperature in this study, was measured using infrared sensor (OMRON Healthcare Co., Ltd., model MC-510-E2, accuracy of 0.1˚C). The sensor is specially designed for babies as it has a smaller head. It provides consistent measurements of the tympanic temperature by recording the maximum value in 10 seconds as established in clinical practice (the tympanic temperature is higher than the rest of the inner ear temperature). Furthermore, in order to improve the accuracy of the obtained value, 2 consecutive measurements were taken and averaged. Infrared sensor (Medscope Ltd., model TH03F, specifically used on "surface mode", accuracy 0.3˚C) was used to measure the skin temperatures. Three consecutive measurements for each skin and clothing location were taken and averaged in order to improve the accuracy of the obtained values. Skin temperature values were very consistent for each set, typically varying only 0.0-0.2 ˚C. Both sensors (OMRON and Medscope) are approved and commonly used in the UK's National Health System (NHS).
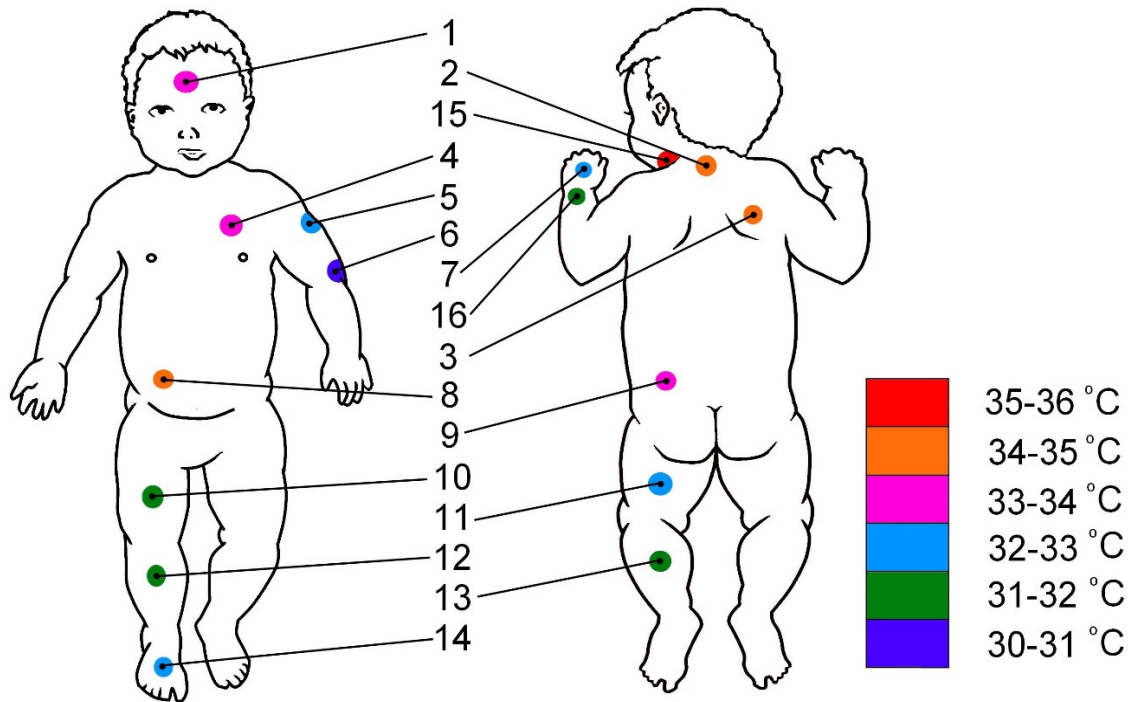
Figure 3: Studied local skin locations and expected temperatures in clothed infants [taken from Garcia-Souto 2016]

3 Data analysis

3.1 Identification of suitable locations

Suitable and non-suitable locations for the monitoring of core temperature were selected according to 3 conditions.

(1) Skin temperature should be characterized by a normal distribution.
(2) Correlation of core and skin temperatures needs to be significant.
(3) In addition to the above it should not depend on many secondary factors to avoid the need for large number of skin temperature threshold values (one per combination of parameters).

The normality of the distribution of core temperature and each local skin temperature, as well as the relevant factors for each of these body temperatures (i.e. conditions 1 and 2) were comprehensively studied and reported by Garcia-Souto and Dabnichki [17].

The relevance of "fever" on each of the local skin temperatures (i.e. condition 3) was investigated and reported in this paper. *Fever* was defined as core temperature over 37.3 ºC and subjects were separated into 2 cohorts according to this Tc_{limit}. The potential 'relationship' between the skin temperature and fever was studied using two methods.

On the one hand, the individual skin temperature distributions were statistically compared to identify if the presence of *fever* was relevant at each given location. The statistical test used in each case was carefully chosen:

- Standard ANOVA test (named *test 1A*): if the whole dataset and each sub-cohort distributions are normal and the variances equivalent (assessed with the Levene test);
- Welch and Brown-Forsythe tests, a robust test of equality of means (named *test 1B*): if the whole dataset and each sub-cohort distributions are normal but the variances are not equivalent (variances assessed with the Levene test);
- Non-parametric tests, i.e. Mann-Whitney U test (named *test 2*): used when normal distributions were not guaranteed, i.e. do not assume normality or equal variances among the cohorts.

On the other hand, Receiver Operator Characteristics (ROC) analysis was performed for the skin temperature at each location following the analysis described by Metz [29]. This was also previously used by Hogan *et al* for fever identification based on forehead temperature [30]. This analysis generates a *sensitivity vs specificity* curve that permits to assess if the studied parameter would be a good marker for a classifying system, and how it would perform using different thresholds. Characterization of the ROC is given, i.e. area under the curve and its standard error (SE), the asymptotic significance, and the lower and upper boundaries for the 95% Confidence Interval (95%CI).

3.2 Characterization of suitable locations

Once the potentially suitable locations for the detection of fever have been identified, we characterize them. Ideally one would be able to measure the local skin temperature of an infant and decide if that value belongs to the skin temperature distributions characteristic of febrile or the no-febrile children, and therefore identify if the child has a fever. Unfortunately this is not straight forward as the two distributions partially overlap causing a degree of uncertainty in the categorization.

In practice, a new skin temperature measurement X_i that falls in the region between the no-febrile (H_0) and febrile (H_a) temperature distributions is compared with a threshold value (k) and categorized accordingly (**Figure 4**). Two standard types of errors could potentially occur: (i) *type I* or α *error*: decide in favour of H_a when H_0 is true hence *creating a false positive*; (ii) *type II* or β *error*: decide in favour of H_0 when H_a is true hence *missing a true positive.* Alpha and beta errors are 'inversely' related, such that the decrease of one causes the increase of the other.

The selection of the threshold value (k) is therefore very important as it defines the alpha and beta errors. For each potentially suitable location a wide range of k values was studied and presented in the alpha-beta curves. These curves were obtained for various sub-cohorts, e.g. based on the infants' age or status, as relevant for the skin location. The optimal k value is obtained by selecting a desirable alpha (or beta) error value and then identifying the

corresponding skin temperature. **Figure 5** features an example with 2 sub-cohorts and a selected alpha error of 20%, leading to beta errors of 65 and 55% and k values of 32.0 and 32.8°C for sub-cohorts 1 and 2 respectively.

Alpha-beta curves contain similar information to the ROCs, with the added benefit that it already presents a measure of the *'sensitivity'* or *'true positive rate'* and *'specificity'* or *'false positive rate'* for all possible values of the threshold k. The main difference is that for the alpha-beta curves the two cohort distributions are first characterized and values obtained looking at the overlap between these two 'normalized' distributions. Instead, the ROCs are directly constructed with the original data set. For completeness, ROC analysis was also included.

Each location was then characterized according to the following features:
- Number of cofounding factors for the local skin temperature
- Mean value of the skin temperature distribution.
- Standard deviation of the skin temperature distribution.
- Average k value at $\alpha = 10\%$ for all the studied sub-cohorts.
- Range of k values at $\alpha = 10\%$, i.e. maximum difference across the studied cohorts.
- Amplitude of the error curve: $k(\alpha \sim 0\%) - k(\alpha = 90\%)$.
- Area under the ROC curve with its standard error (SE)
- Lower and upper boundaries for the 95%CI for areas under the ROC.

A rank of the more suitable locations was derived attending to the following criteria:
- Statistical correlation with core temperature (must).
- Small number of factors for the skin temperature value at that location (+1 point).
- High amplitude in the error curves was preferred as a small change of k value will not affect the percentage of error so strongly (+ 1 point).
- Small range of k values among the studied sub-cohorts was preferred as then the average k value could be used without generating substantial α and β errors (+ 1 point).
- High area under the ROC curve as average (+1 point) and within the 95%CI (+1 point).
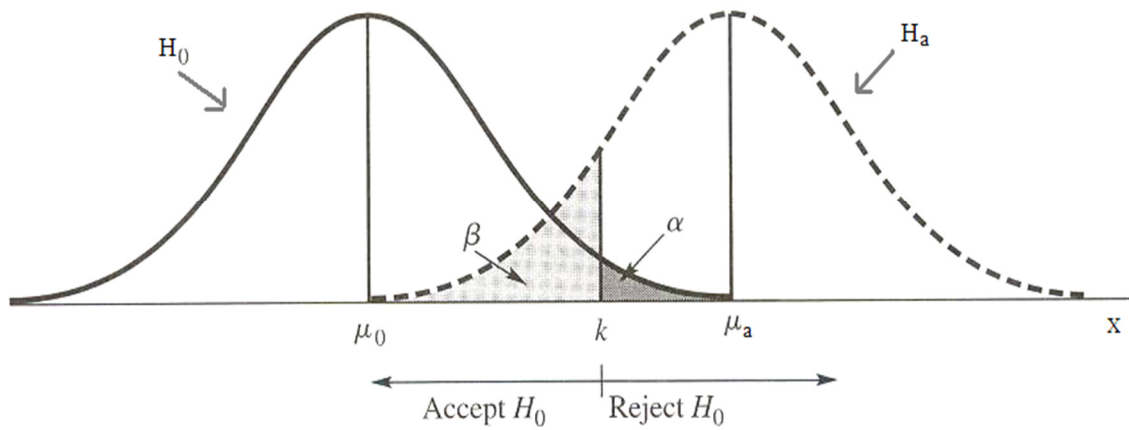
Figure 4: Representation of 2 overlapping distributions and corresponding alpha and beta errors for a given threshold k.
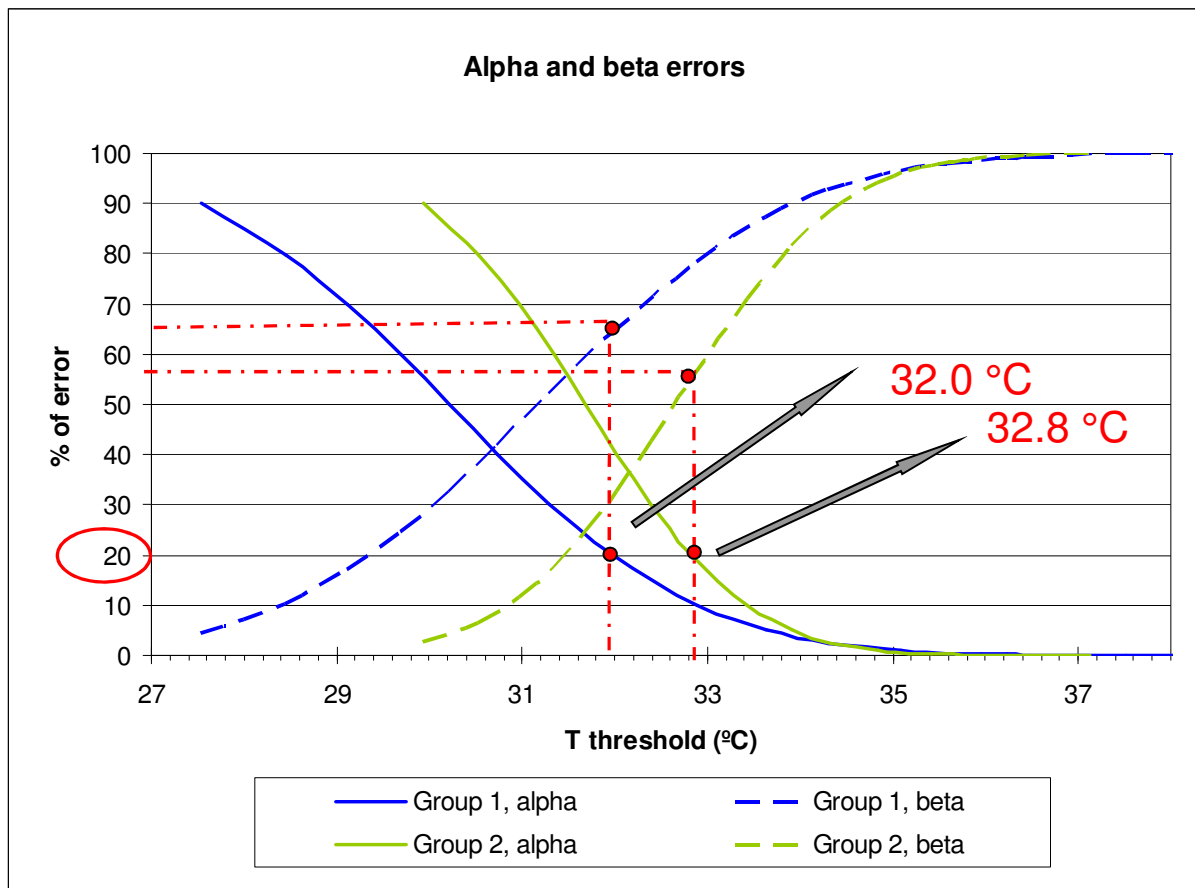
Figure 5: Example of k value selection on a skin location with 2 different studied subgroups of babies

3.3 Sensor/system design and characterization

Various monitoring system designs were studied, those were: (i) a single sensor at various skin locations; and (ii) various multi-sensor system designs, within a (a) single or (b) multiple locations. These designs assume a basic binary sensor located over the skin. The sensor is ON

when the skin temperature is greater than a threshold value (*k*) indicating high probability of "fever", and it is OFF when the skin temperature is below the threshold, indicating high probability of "no fever". The four possible cases are presented in **Table 1**, with the associated probability of occurrence. Alpha and beta errors are expressed in terms of percentages, while p_F is the probability of a random child having fever also known as *prevalence*, and therefore *1- p_F* is the probability of the child being healthy (no-fever).

	Children with fever p_F	Children with no fever $1- p_F$
Sensor ON	$p_F . (1 - \beta)$	$(1-p_F) . \alpha$
Sensor OFF	$p_F . \beta$	$(1-p_F) . (1 - \alpha)$

Table 1: Possible cases of fever assessment by a single binary sensor. Values normalized to the total number of cases, hence addition of all values in the table is equals to 1.

3.3.1 Single sensor

The simplest design option is to use of a single sensor in one of the appropriate skin areas. The most suitable locations were studied. The characterization of the sensors (as described below) was obtained for two fixed alpha error values, 10 and 20%.

The total error (ε_T) associated to a single fever sensor is given by **equation 1**, which is based on **Table 1**. The probability that the sensor correctly identifies fever (*p*) is given by *1-ε_T*, and this corresponds to the accuracy (**equation 2**). The probability of a child having fever when the sensor is ON and when it is OFF is given by **equation 3 and 4** respectively. ROC analysis was also used to calculate the accuracy, defined as the number of correct assessments divided by the total number of cases (**equation 5**); this is based on the raw data. This was performed for completeness and to allow for comparison.

$$Total\ error = \varepsilon_T = p_F.\beta + (1- p_F).\alpha \qquad (1)$$

$$Accuracy_{based\ on\ \alpha-\beta curves} = 1-\alpha + p_F.(\alpha - \beta) = 1 - \varepsilon_T \qquad (2)$$

$$p_i(fever\ when\ sensor\ is\ ON) = \frac{p_F(1-\beta_i)}{p_F(1-\beta_i)+(1-p_F)\alpha_i} \qquad (3)$$

$$p_i(fever\ when\ sensor\ is\ OFF) = \frac{p_F\beta_i}{p_F\beta_i + (1-p_F)(1-\alpha_i)} \qquad (4)$$

$$Accuracy_{ROC} = sensitivity * p_F + specificity * (1- p_F) \qquad (5)$$

$$where\ sensitivity = 1 - \beta$$
$$specificity = 1 - \alpha$$

### 3.3.2 Multi-sensor system: single sensors at various locations

Two system designs formed by 3 and 5 sensors located at various places were studied, aiming to increase the probability of correctly identifying fever in children. Odd numbers of sensors were used to aid the interpretation of the readings, e.g. the sets of 3 or 5 sensors would indicate high probability of fever if 2+ or 3+ sensors are ON respectively. The probability that the system gives a correct assessment of fever (p_T) for a 3-sensors and a 5-sensors set is given by the **equations 6 and 7** respectively, based on the p values for individual sensors. Equivalent equations were used to estimate the total alpha and beta errors.

$$p_{\text{3-sensors-set}} = \prod_{i=1}^{3} p_i + \sum_{\substack{k=1 \\ i \neq j \neq k \\ i,j \in \{1,2,3\}}}^{3} p_i p_j (1 - p_k) \tag{6}$$

$$p_{\text{5-sensors-set}} = \prod_{i=1}^{5} p_i + \sum_{\substack{k=5 \\ i \neq j \neq t \neq k \neq z \\ i,j,t,k,z \in \\ \{1,2,3,4,5\}}}^{1} \sum_{z=k-1}^{1} p_i p_j p_t (1 - p_k)(1 - p_z) \tag{7}$$

3.3.3 Multi-sensor system: multiple sensors at single location

A system design formed by 5 sensors located at a single skin area was studied, aiming to increase system's reliability and adaptation to user. Each individual sensor has a different temperature threshold value, allowing for different levels of alert to which the user chooses to respond based on experience. Sensors are placed in a line (forming an array), ordered based on their temperature threshold with the lowest at the bottom. There are only 6 possible readings (see **Figure 6**) as in all cases when sensor i is ON, sensors j with $j<i$ are necessarily ON as well. The probability of a child having fever when the i-th sensor is ON was calculated by **equation 3**. Thresholds and accuracy of each sensor were identified by means of the alpha-beta curves for the specific location, and selected to cover a wide range.
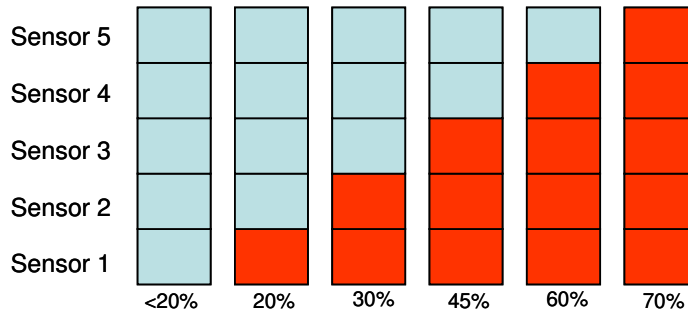
Figure 6: Six possible readings of a 5-sensors set (blue = sensor is OFF; red = sensor is ON)

4 Results

4.1 Relevance of fever on local skin temperatures

Characterization of local skin temperature distributions for babies with core temperature under and over 37.3ºC (threshold defining *fever*) is given in **Table 2**. As example, **Figure 7** shows the two normalized skin temperature distributions at *T5*. Results of the statistical analysis investigating the relevance of *fever* is given in **Table 3** for each location, along with information about their specific ROC that indicates how suitable the location is for fever identification. Both analyses gave the same results. As example, **Figure 8** shows the ROC curve corresponding to local skin temperature T5 as individual indicators of 'fever'. All ROC curves can be seen in **Appendix A**.

	Tc < 37.3°C			Tc >= 37.3°C		
	N	**Mean**	**SD**	**N**	**Mean**	**SD**
T1	518	33.76	0.72	21	34.74	0.85
T2	517	34.51	0.87	21	35.1	1.3
T3	517	34.55	0.83	21	35.50	0.74
T4	510	33.95	0.98	21	34.1	1.5
T5	501	32.80	0.98	21	33.9	1.1
T6	230	31.1	1.3	10	31.3	1.3
T7	516	32.0	1.8	21	32.9	1.9
T8	509	34.48	0.91	21	35.1	1.2
T9	508	33.9	1.0	21	34.9	1.0
T10	510	31.7	1.4	21	32.3	1.5
T11	510	32.5	1.2	21	33.3	1.1
T12	502	31.6	1.4	21	32.5	1.5
T13	500	31.3	1.5	21	31.6	1.4
T14	345	33.0	1.8	16	33.4	1.9
T15	408	35.13	0.77	17	35.65	0.56
T16	415	31.0	1.5	17	31.7	1.6

Table 2: Characterization of local skin temperature distributions for babies with T_c under and over 37.3ºC. (Values for locations T5-6 and T12-14 only represent the cases where the location was covered with clothing)

Temperature distribution on the right arm in upper location for two ranges of core temperature

(a)	**(b)**
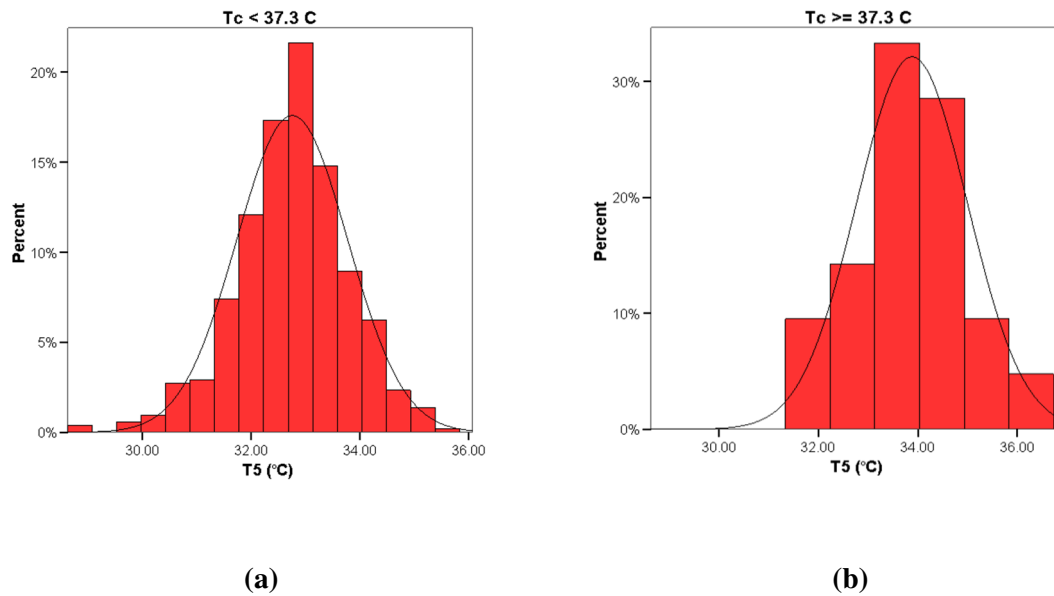
Figure 7: Normalized temperature distribution on the right arm in upper location for two cohorts. a) *Tc< 37.3°C.* b) *Tc≥ 37.3°C.*

Location	Fever	ROC Area (SE)	ROC Asymptotic Sig.	95%CI boundaries
T1	**1A**	0.817 (0.058)	**0.000**	0.70 - 0.93
T2	**2**	0.727 (0.065)	**0.000**	0.60 -0.85
T3	**1A**	0.804 (0.048)	**0.000**	0.71 -0.90
T4	1B	0.548 (0.080)	0.454	0.39 -0.71
T5	**1A**	0.766 (0.059)	**0.000**	0.65 -0.88
T6	1A	0.523 (0.094)	0.802	0.34 -0.71
T7	**1A**	0.630 (0.064)	**0.043**	0.50 -0.76
T8	**1A**	0.656 (0.069)	**0.015**	0.52 -0.79
T9	**1A**	0.770 (0.057)	**0.000**	0.66 -0.88
T10	**1A**	0.615 (0.061)	**0.073**	0.50 -0.73
T11	**1A**	0.662 (0.056)	**0.012**	0.55 -0.77
T12	2	0.605 (0.063)	0.132	0.48 -0.73
T13	1A	0.523 (0.064)	0.737	0.40 -0.65
T14	1A	0.574 (0.074)	0.316	0.43 -0.72
T15	**2**	0.702 (0.063)	**0.005**	0.58 -0.83
T16	**2**	0.670 (0.064)	**0.018**	0.54 -0.79

Table 3: Summary of the local skin temperature dependency on the presence of *fever* and its capacity for fever identification according to ROC analysis. Statistical tests used were: (1A) ANOVA with homogeneous variance, (1B) ANOVA with different variances, (2) Mann-Whitney U test. Results are: (Black) Distributions are equivalent, and (Red) Distributions are statistically different hence *fever* is relevant. (Values for locations T5-6 and T12-14 only represent the cases where the location was covered with clothing)
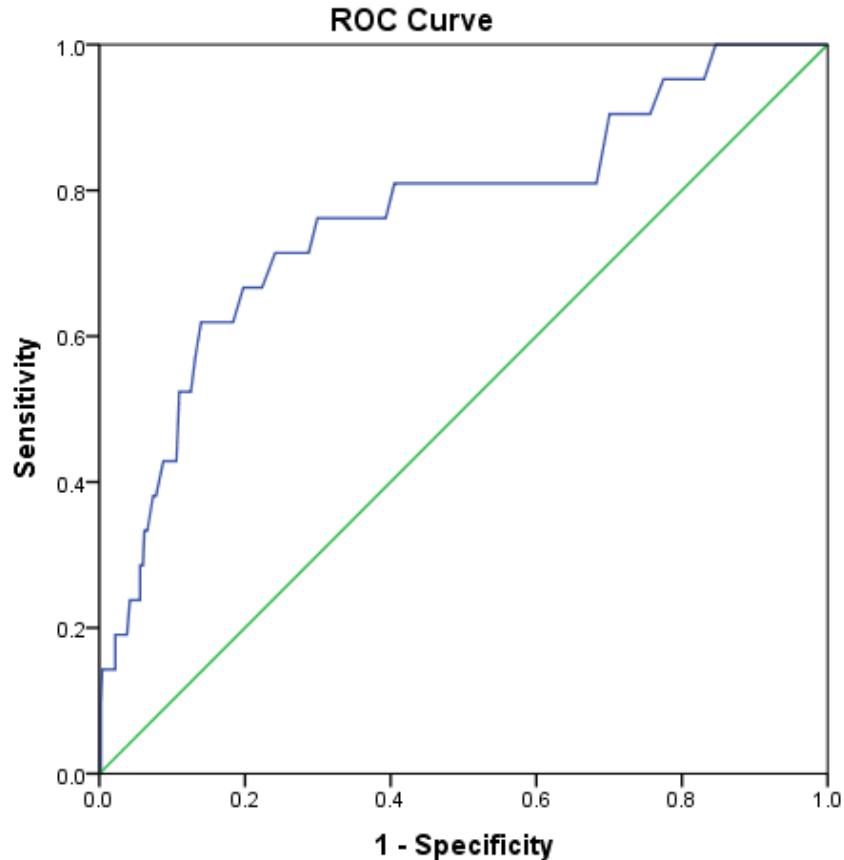
Figure 8: ROC curves for local skin temperature T5 as predictors of 'fever'.

4.2 Non suitable locations

Several locations had to be withdrawn as they do not satisfy one or several of the required conditions for serving as a good estimator of core temperature, as defined in **section 3.1**. Locations *T2* and *T15* (back of the neck and lateral) had temperature distributions more pointed than normal distribution [17]. Skin temperature at *T10, T12* and *T14* depend on a high number of factors [17], making them impractical for the estimation of core temperature. Skin temperature at locations *T4, T6, T12, T13* and *T14* (left upper chest, left arm, right shin, left calf and right instep) were not significantly different with the presence of *fever* (**Table 3**).

4.3 Characterization of suitable locations

Alpha and beta curves were obtained for those potentially suitable locations. As example, the specific case of T5 is given in **Figure 9**; curves for all the locations can be seen in **Appendix B**. These locations were then characterized and ranked as described in section 3.2 (**Table 4**). Only samples of temperature in covered skin have been taken into account for T5. All others were typically covered or exposed.
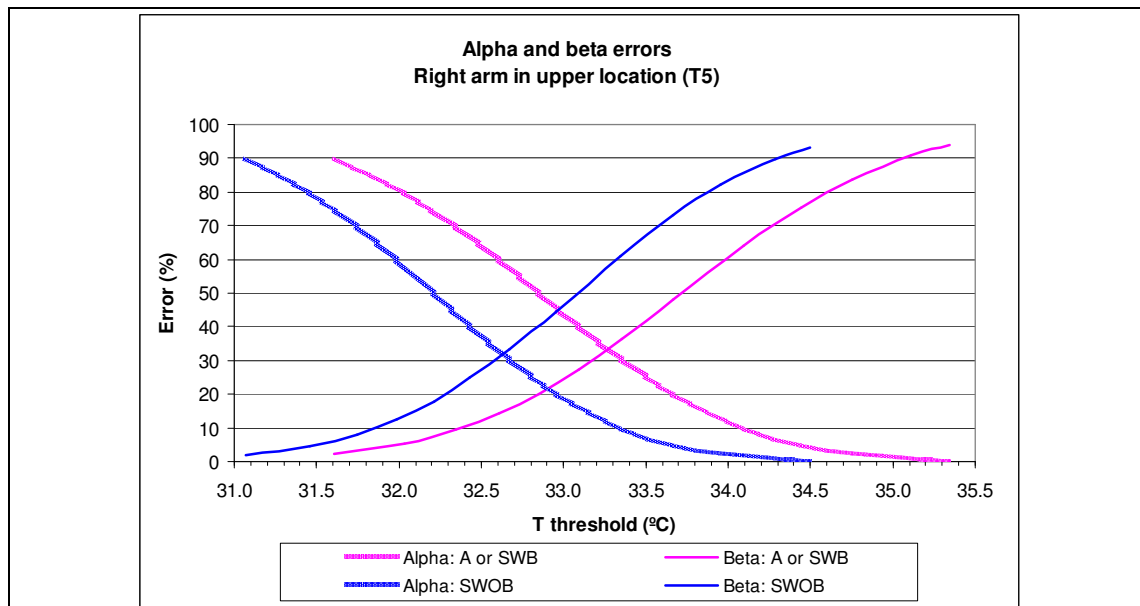
Figure 9: Alpha (detected false positives) and beta (missed true positives) errors for different skin temperature thresholds at right arm in upper location. Nomenclature: (*A)* awake; (*SWB*) sleeping with a blanket; (*SWOB*) sleeping without a blanket.

	T1	T3	T5	T7	T8	T9	T11	T16
N sub-cohorts	6	**2**	**2**	9	**2**	6	**4**	9
Mean temperature (°C)	33.8	34.6	32.8	32.0	34.5	33.9	32.6	31.1
Standard deviation (°C)	0.75	0.84	1.05	1.83	0.93	1.02	1.24	1.48
Temperature at =10% (°C)	34.3	35.0	33.7	34.2	35.8	35.5	34.3	33.3
Amplitude of error curve (°C)	3.0	3.1	**4.0**	**8.0**	3.3	**4.6**	**4.7**	**6.7**
Range of T limit values at =10% (°C)	1.1	**0.7**	**0.8**	2.8	**0.4**	1.5	1.1	2.1
ROC Area (and stand. Error)	**0.82 (0.06)**	**0.80 (0.05)**	**0.77 (0.06)**	0.63 (0.06)	0.66 (0.02)	**0.77 (0.06)**	0.66 (0.06)	0.67 (0.06)
95%CI boundaries	**0.70 - 0.93**	**0.71 - 0.90**	**0.65 - 0.88**	0.50 - 0.76	0.52 - 0.79	**0.66 - 0.88**	0.55 - 0.77	0.54 - 0.79
Added points	2	4	5	1	2	3	2	1
Ranking	4	2	1	5	4	3	4	5

Table 4: Characterization of suitable locations for core temperature monitoring attending to temperature distribution, alpha and beta error curves, and ROC curves. (**Bold** represents good characteristics, counting towards the added points)

4.4 Examples of sensor/system design and characterization

We studied the efficiency of several sensors/system designs for the correct identification of fever in children. These are intended to be used for any child regardless of age (provided it is up to 2 years old), gender, BMI or activity level. The only restriction applied is that all data used corresponds to local skin measurements taken when the location was covered.

Choosing to design a *'generic'* system was a practical decision taken at the expense of providing a more accurate system for a particular cohort. For example, one could design a system adapted to awake children aged between 8 and 16 months and with BMI in the range of 16.53 to 19.25 kg/m^2 corresponding to the *normal healthy* category, or any other group of children for that matter. However a generic system provides a wider usability. Benefits and limitations will be discussed later on.

4.4.1 Single sensor

From the ranking of suitable locations *T3*, *T5*, *T8*, *T9, T11* and *T16* were selected for further evaluation for two fixed values of alpha error, those were 10 and 20% (**Table 5**) following the procedure described in 3.3.1. *T1* and *T7* were not selected as they are typically exposed areas. The probability of a random child having fever (*p_F or prevalence*) was taken as 0.15.

Accuracy values are very similar across the different sensors, and consistently higher when calculated with the ROC analysis. Probability of the sensor being ON when the child has a fever is more indicative in this case, discriminating among the locations. The best single sensors seem to be T3, T5, and T9 for either of the alpha values, and T8 when alpha is 0.1. Corresponding thresholds can be seen in **Table 5**.

α	Location	β	$ε_T$	Accuracy	p(ON+ fever)	p(OFF+ fever)	ROC based accuracy	Threshold k (°C)
0.1	T3	0.56	0.17	0.83	0.44	0.10	0.92	33.4
	T5	0.56	0.17	0.83	0.44	0.10	0.92	31.6
	T8	0.69	0.19	0.81	0.35	0.12	0.91	33.3
	T9	0.61	0.18	0.82	0.41	0.11	0.92	32.5
	T11	0.77	0.20	0.80	0.29	0.13	0.91	30.8
	T16	0.77	0.20	0.80	0.29	0.13	0.91	29.2
0.2	T3	0.37	0.23	0.77	0.36	0.075	0.83	34.0
	T5	0.41	0.23	0.77	0.34	0.083	0.82	32.1
	T8	0.56	0.25	0.75	0.28	0.11	0.82	33.8
	T9	0.44	0.24	0.76	0.33	0.088	0.82	33.2
	T11	0.61	0.26	0.74	0.26	0.12	0.82	31.5
	T16	0.63	0.26	0.74	0.25	0.12	0.81	30.0

Table 5: Characterization of single fever sensors

4.4.2 Multiple-sensor system: single sensors at various locations

A 3-sensors set and a 5-sensors set were studied for two fixed values of single-sensors alpha error (10 and 20%) following the procedure described in 3.3.2 (**Table 6**), i.e. defining 2+ or 3+ sensors ON as a sign of fever. The best performing individual sensors were selected (T3, T5, T9 and T8 in that order). T11 and T16 are almost identical in terms of performance. T11 was selected for the 5-sensors system so it covers both areas on the upper body and lower body. Alternative, T16 could be selected so all the sensors can be placed in an upper-body garment and the performance of the system would be very similar.

Location	$\alpha_i = 0.1$					$\alpha_i = 0.2$				
	p_T(right)	α	β	p(ON+ fever)	p(OFF+ fever)	p_T(right)	α	β	p(ON+ fever)	p(OFF+ fever)
T3-T5-T9	0.92	0.028	0.61	0.71	0.10	0.86	0.10	0.36	0.52	0.066
T3-T5-T9-T8-T16	0.96	0.013	0.75	**0.77**	**0.12**	0.90	0.058	0.47	0.61	0.080

Table 6: Performance characterization for systems of several fever sensors

4.4.3 Multiple-sensor system: multiple sensors at single location

The use of 5 sensors on the upper arm (T5) was investigated following **section 3.3.3**. Characteristic values for those sensors are presented in **Table 7**. A 3-set sensors could also be used by choosing 3 out of the 5 presented sensors. This location was selected for its convenience, while the performance is not compromised according to **Table 5**.

Sensor	S1	S2	S3	S4	S5
$T_{threshold}$	32.6	33.4	34.1	34.8	35.3
α	0.60	0.30	0.1	0.025	0.005
β	0.12	0.31	0.56	0.77	0.90
p(ON,fever)	0.21	0.29	0.44	0.62	0.78
p display	20%	30%	45%	60%	80%

Table 7: Characterization of array of sensors located at upper arm (T5)

5 Discussion

5.1 What makes a skin location to be good for fever occurrence monitoring?

First of all, the skin temperature needs to be significantly related with the core temperature. Various locations were found where the skin temperature significantly increases with the occurrence of fever (defined by a 37.3°C threshold). This dependency is observed mainly at the trunk locations (except upper chest), neck and forehead, and at some limb locations (hands, upper arm and thigh). However it is not observed at the most peripheral areas of the lower limbs (T12, T13 and T14) or the lower arm (T6). These observations are to be expected as, due to thermoregulation mechanism, the temperature at central locations is kept more stable and close to the core temperature, while those at the periphery of the body vary more significantly with the room temperature [16]. The current study was conducted at mild room temperatures, which explains why this effect seems to be restricted only to the most peripheral areas of the legs. A small overlap between the local skin temperature distribution for *febrile* and *non-febrile* cohorts is also desirable and would make the specific skin area better suited for fever recognition. This translates into broader *alpha* and *beta error* curves for the specific location, larger areas under the ROC curves and smaller error values for a given *k* threshold.

Besides, in a good location the skin temperature should not depend on many other personal factors other than *'fever'*, such as *age*, *status* or *BMI*, otherwise it is difficult to set the baseline temperature for healthy children and corresponding thresholds for *fever* detection. The factors for each skin temperature for toddlers were comprehensively reported previously [17] and are used in the current work.

Finally, having skin temperature distributions that are close to a normal distribution also makes the probabilistic analysis simpler (with the alpha and beta curves), although this is not strictly necessary. This was used as one of our selection criteria so a better estimate of the system performance could be calculated.

5.2 System design

There are three main aspects to consider in the system design.

a) Type of error that is most critical/suitable, with options being:
   i. Minimize alpha errors: More children with a fever will not be detected by the system but the number of false alarms will be reduced.
   ii. Minimize beta errors: Most of the children with a fever will be detected by the system although it will also diagnose many healthy children as having a fever.
   iii. Compromise: Choose a limit temperature which corresponds with no extreme values for both alpha and beta errors.
b) Type of sensors to be used in the system:
   i. Generic sensors and thresholds that relate to children of a range of characteristics (age, BMI): The performance will vary from group to group but it simplifies the 'product selection' and production cost.
   ii. Specific sensors for each group: Sensors will perform better as far as it is used in the cohort for which they were adapted.
c) Complexity of the system: One-sensor system vs. multiple-sensor system, which affects performance, robustness, and cost.

The selection of the *k* value (or threshold temperature) is highly dependent of the system aim and its design criteria point (a) and (b). This paper not only presents some examples, but also provides relevant information (full alpha-beta curves and ROC curves) for other designers to set their own criteria and design the fever detection system accordingly. Various cases of one-sensor and multiple-sensor system (design criteria (c)) are also reported and discussed.

5.2.1 Selection of *k* value

According to design aspect (a), *k* can be chosen to minimize either the percentage of alpha or beta errors, or alternatively used the common medical approach of finding a compromise between the two. Alpha and beta errors are expressed in terms of percentages, but their calculation is based on different groups of children, i.e. **healthy children** and **children with a fever** respectively. Therefore at equal values of alpha and beta errors, alpha deems a larger number of 'mistakes' in normal circumstances as the prevalence of fever is low. This paper considers that a reasonable assumption is that a fever identification device should minimize

the alpha error, i.e. reducing the occurrences where the childminders are falsely alerted, or otherwise the childminders would stop using the device. Childminders would trust more a device that gives a right diagnosis of a fever despite it overlooking some fever cases. Hence, the imposition of a maximum alpha error of 10-20% was considered appropriate.

On the other hand, according to design aspect (b), a general k value or threshold can be selected for a given location (based on the average alpha-beta curves or generic ROC curves), or rather given for individual sub-cohorts to be more precise, reflecting the dependencies of the skin temperature distribution at that location. However, here the increase on the system accuracy has the downside of making the product more complex to use, as a specific 'product' needs to be matched with a specific type of child or otherwise it would not deem relevant assessment of the core temperature, and/or would not perform as expected. This makes parents and carers less likely to use the product confidently, and with time it falls into disuse. The use of various k values depending on the children cohort also makes the product more expensive, as various types of sensors need to be produced. The sensors also need to be matched with specific children's clothes, increasing the amount of stock needed. One could argue that as the growth of the children is fast at those ages, it is feasible to have different k values depending of age. Greater variety of garments increases the cost but also potentially increases the number of sales. In conclusion, there is not right or wrong answer. The design criteria would typically define the required system accuracy and simplicity, and therefore will dictate which option should be used.

In the examples reported in this paper, one single k value per location is used for all children with the only condition that they are up to 2 years old, i.e. the scope of this study. The age, gender, BMI and activity of the child, e.g. if s/he is sleeping or awake, are not noted. This provides a practical option in terms of easiness-to-use and maintaining cost production low. Because the system is generic, it will perform better for certain groups than what it was reported.

5.2.2 Selection of the system complexity

The use of one single sensor was compared with the use of a multi-sensor system, either (i) distributed across various body locations or (ii) at the same location but with different k-thresholds. Garments with single sensors are simpler and more economical to produce; cost might increase proportionally to the number of sensors. However, multi-sensor systems of either type offer an increase on reliability and allow for some degree of personalization based on the carer's observations of the system response to a specific child. They are therefore recommended.

Individual generic sensors (from **Table 5**) have accuracy of 80-83 or 91-92% at α=10% depending of the analysis method (alpha-beta curves or ROC curves), but with low probability of being ON during fever (ranging 29-44%). Sensors adapted to a specific group of children would perform better. However, the use of a combination of sensors increases the detection performance. One example of (i) is the set of generic sensors located at T3-T5-T9-T8-T16. This combination has an accuracy of 96% and increases the probability of the

system being ON when the child has a fever by a factor of ~2with respect to a single average sensor (up to 77%), while still keeping the probability of missing a 'fever' case low (12%). This is highly advisable to promote long-term use of the system and trust by the carers. It also allows the carers to identify which of these locations might be more reliable for their child, hence improving the efficiency of the system further. A 3-sensor set also performs well (accuracy = 92%, probability of the sensor being ON during fever = 71%, sensor OFF during fever = 10%) and it would be more economical.

One example of (ii) is the 5-sensor-set at T5. Here, each sensor requires a higher skin temperature than the previous ones to get activated, but also represents a higher probability of fever occurrence. Therefore, such system provides various degrees of fever alarm, depending on the number of sensors that become ON. Based on experience, the carer might be able to identify the most suitable alarm threshold for a given child. For example, a parent would disregard the system if only 1, 2 or 3 sensors are ON out of the available 5 sensors, but would consider a real alarm if 4 or 5 were activated.

5.3 System limitations

Local skin temperature inherently depends on a range of factors, and present large variability even for seemingly similar cases. This will always remain a limiting factor for the accuracy of any fever detection system based on the skin temperature readings. However, such systems are still valuable as it provides an indication of potential fever, therefore triggering a follow up by a more traditional core temperature method.

Due to the young age of the subjects, the present study was only allowed to be undertaken under a 'narrow' range of room temperatures as set up in the nurseries. Large variations in the room temperature will have an effect in the skin temperatures and therefore the efficiency of the system, and might be studied in the future if ethical approval can be obtained. However, recommended locations are not at the most peripheral areas of the limbs, which helps the skin temperatures to be more independent of the environmental temperature, reducing the impact of this limitation.

Similarly, working with this subject group strongly limited the measurement techniques that we could use to ascertain the core temperature, with typical 'gold standards' being not allowed. Following the recommendations within the *"Fever in under 5s: assessment and initial management – NICE Clinical Guideline 160"* [1] and practical reasons, an IR ear thermometer was used as to obtain an approximate core temperature value. This means that the system has been designed to identify an elevated tympanic temperature rather than core temperature, but it is still suitable to screen for fever. Also, even if a 'gold standard' was used to obtained more accurate core temperature values to inform the development, the fever detection systems would probably have improved just marginally, as local skin temperature variability presents the largest source of error.

Only children with marginal fever (<38.1ºC) were present at the nurseries. This leads to an overestimation of the beta errors, which would be advantageously lower in a not biased group of children, as individuals with high fever are more likely to be detected.

In any case the sensor should be embedded in the babies clothing and not in an extra band or garment, as this would change the heat transfer at that particular skin location and it would induce to error [31] unless the extra insolation is compensated.

5.4 Future work

The priority for future work would be to acquire further skin temperature data in children with higher fever than those present in the nurseries. This would probably need to be performed at hospitals, walk-in-centres or general health practices. This would allow us to reassess the temperature distributions, and more accurately estimate the system performance. Most likely, the k values to detect fever would increase, therefore decreasing the probability of false positives even further. In addition, it might allow for the locations at the neck (T2 and T15) to be used; these were discarded because they had a large Kurtosis but were otherwise potentially good indicators of the core temperature.

The next step would be to run field tests with the developed system and experimentally establish the system performance.

The methodology presented in this paper is applicable to detect fever regardless the temperature sensor used. Hence, another area to explore could be the combination of this methodology with temperature sensors capable of communicating with the Internet of Things or smart phones, allowing for remote fever monitoring. This would also allow to adapt to a given child in terms of age, gender, BMI, etc., perhaps via a mobile app.

6 Conclusions

Monitoring systems for the identification of fever in young children is highly desirable. Clinical accuracy is not required, but they should be non-intrusive and practical for everyday life, at home, nurseries, etc. Also, these systems need to be built on data collected from toddlers rather than adults, as they present different physiological characteristics. Several system designs for the continuous non-invasive identification of fever based on skin temperature measurements were studied, various single sensor and multi-sensor systems. These were designed for a generic child regardless gender, BMI, activity or age as long as they are below 2 years old. Identical analysis can be made for any group of children, for example something adapted to the particular case of awake children aged between 8 and 16 months and with BMI in the range of 16.53 to 19.25 kg/m^2. In that regard, the current paper provides all the necessary data so other members of the community can design their own sensors and system based on their own criteria.

For the studied cohort, single sensor systems have a probability of positive identification of fever on feverish children of ~30-44%, using an α (detected false positives) of 10%. Multi-sensor systems of sensors distributed among different locations increase the probability of positive identification of fever to 71% (set of 3 sensors) and 77% (set of 5 sensors) without significantly increasing the probability of false positives. The use of several sensors of different characteristics at the same location is also beneficial, as it displays different levels of probability of the child having fever, up to 80%. It is important to note that this type of systems will probably detect earlier than a human observer when the temperature rises sufficiently; so in this sense even if not completely successful, could still deliver positive results.

We conclude that the non-intrusive and continuous monitoring of fever in children under 2 years is possible by using the technology described in this study. This is of great importance as it would allow the detection of a greater number of feverish cases in early stages without having to measure the core temperature continuously. This might also help to detect underlying infections that can be life-threatening, and perhaps avoid spread among other children, e.g. in schools or nurseries. The fever monitoring system we would recommend is the multi-sensor system with a set of 5 sensors distributed among locations $T3$, $T5$, $T8$, $T9$ and $T16$, each sensor with an independent $\alpha_i = 10\%$. For this system, the child is likely to have fever in 77% of the cases in which the system alerts the childminder; and 12% of the children will have fever when the system is OFF. Missing 12% of the children with fever is reasonably taking in consideration that this is not a clinical system but an indicator. This wearable system provides a non-invasive solution to continuously monitoring for fever; hence it provides an early alert system for the career. This is particularly useful when the ratio children-to-careers is high such as in nurseries. However, when carers suspect that the child have a fever either because this system has alert them or otherwise, more traditional core temperature measurement should be used for confirmation such as an ear thermometer.
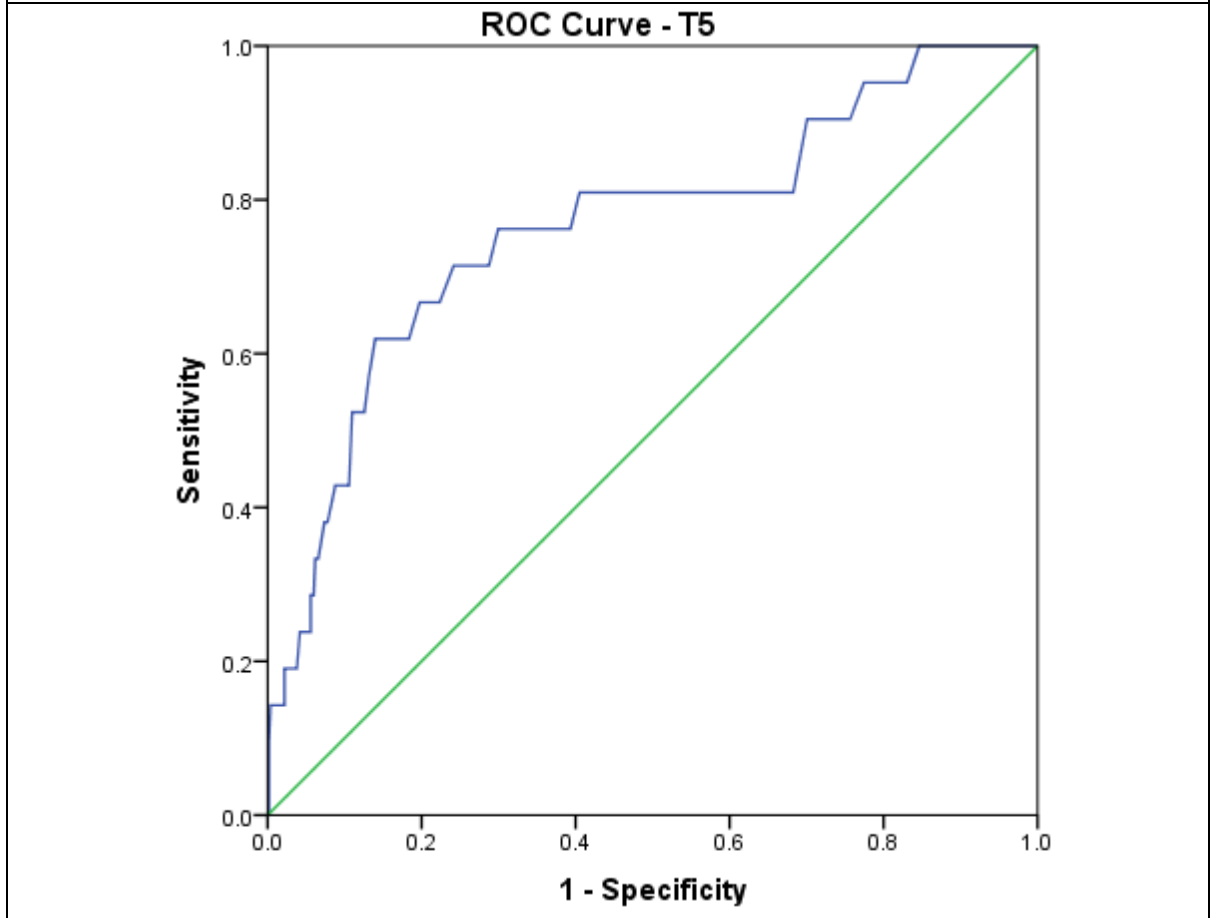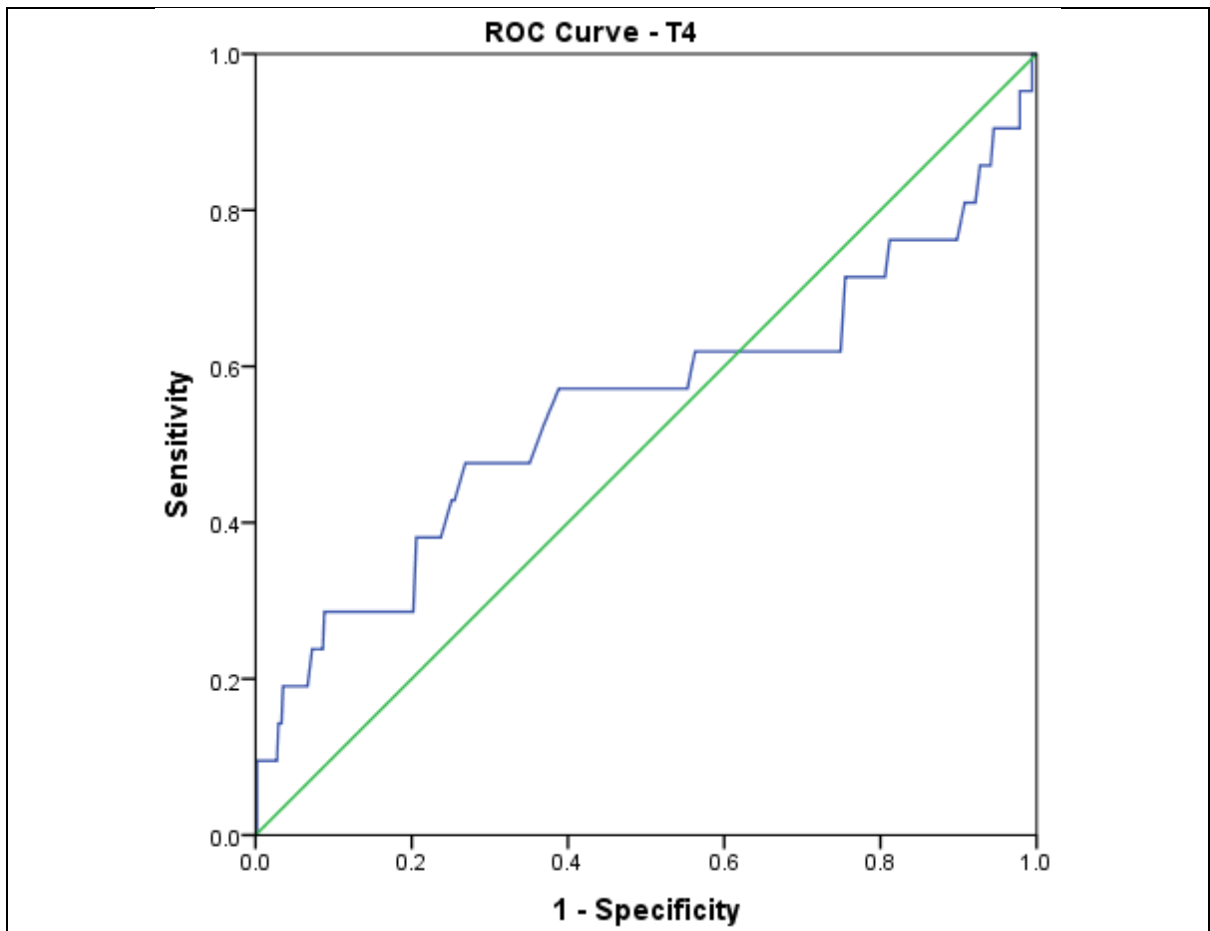
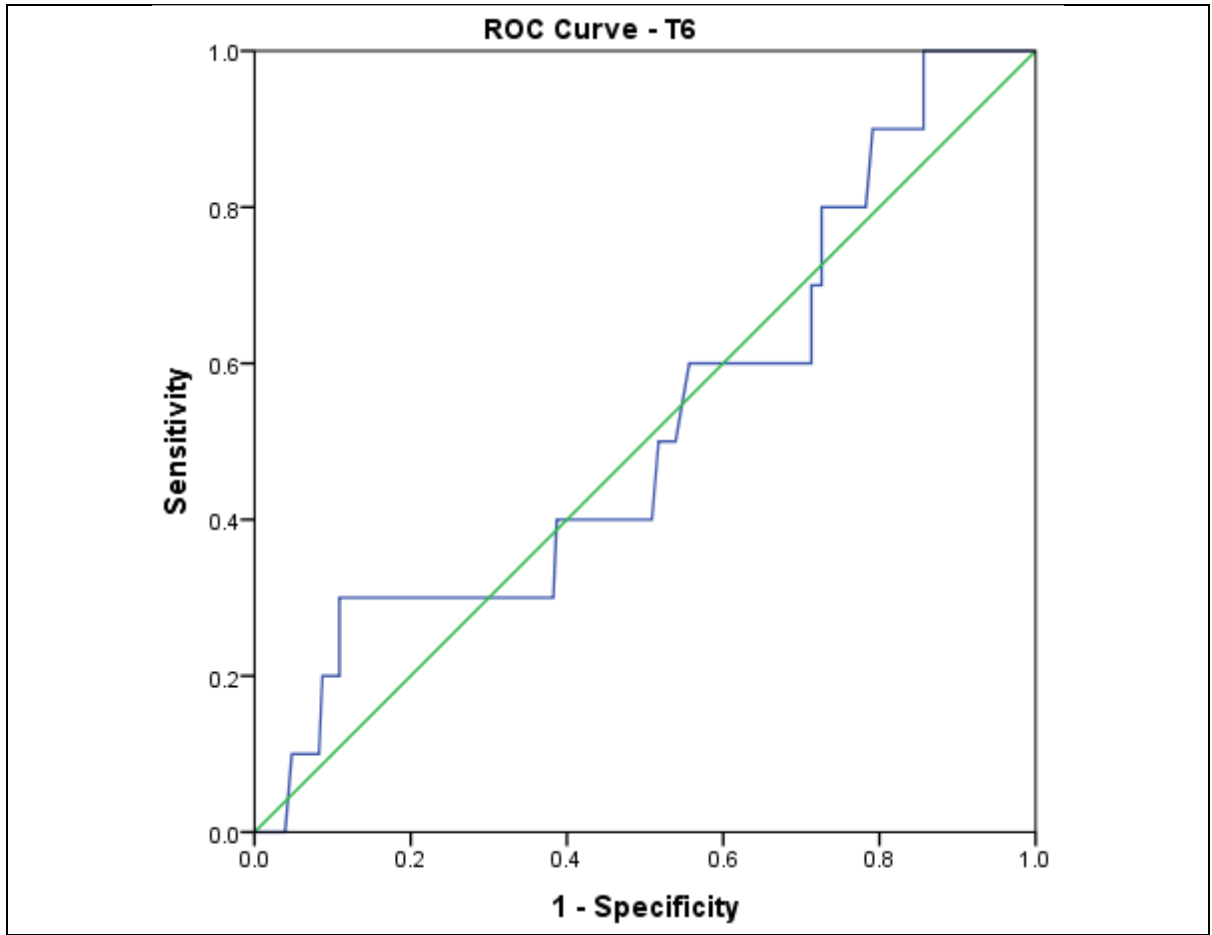## Acknowledgements

Appendix A

Receiver Operator Characteristics (ROC) curves obtained for all the potentially suitable locations for fever monitoring and detection (**Figure A1**).
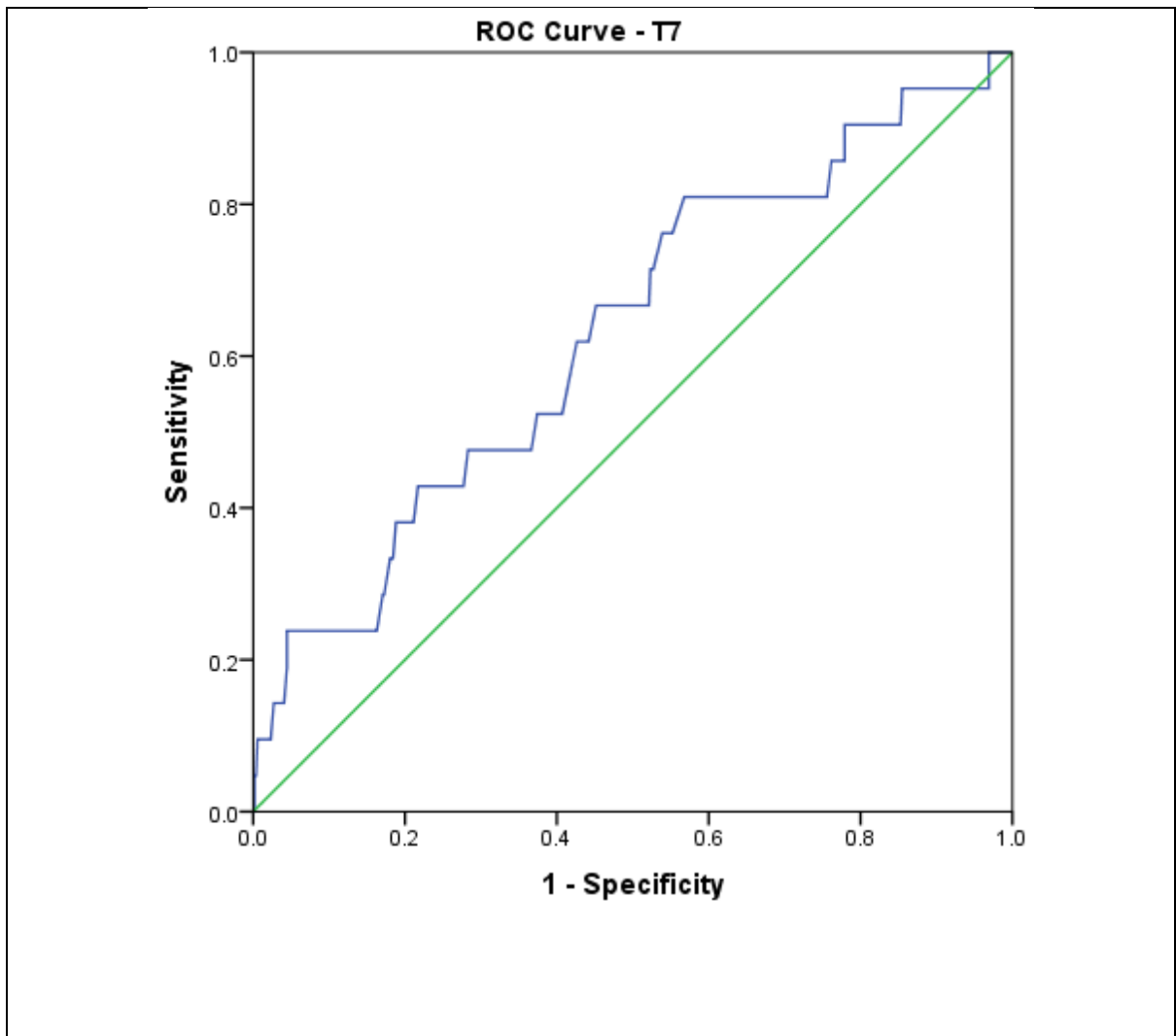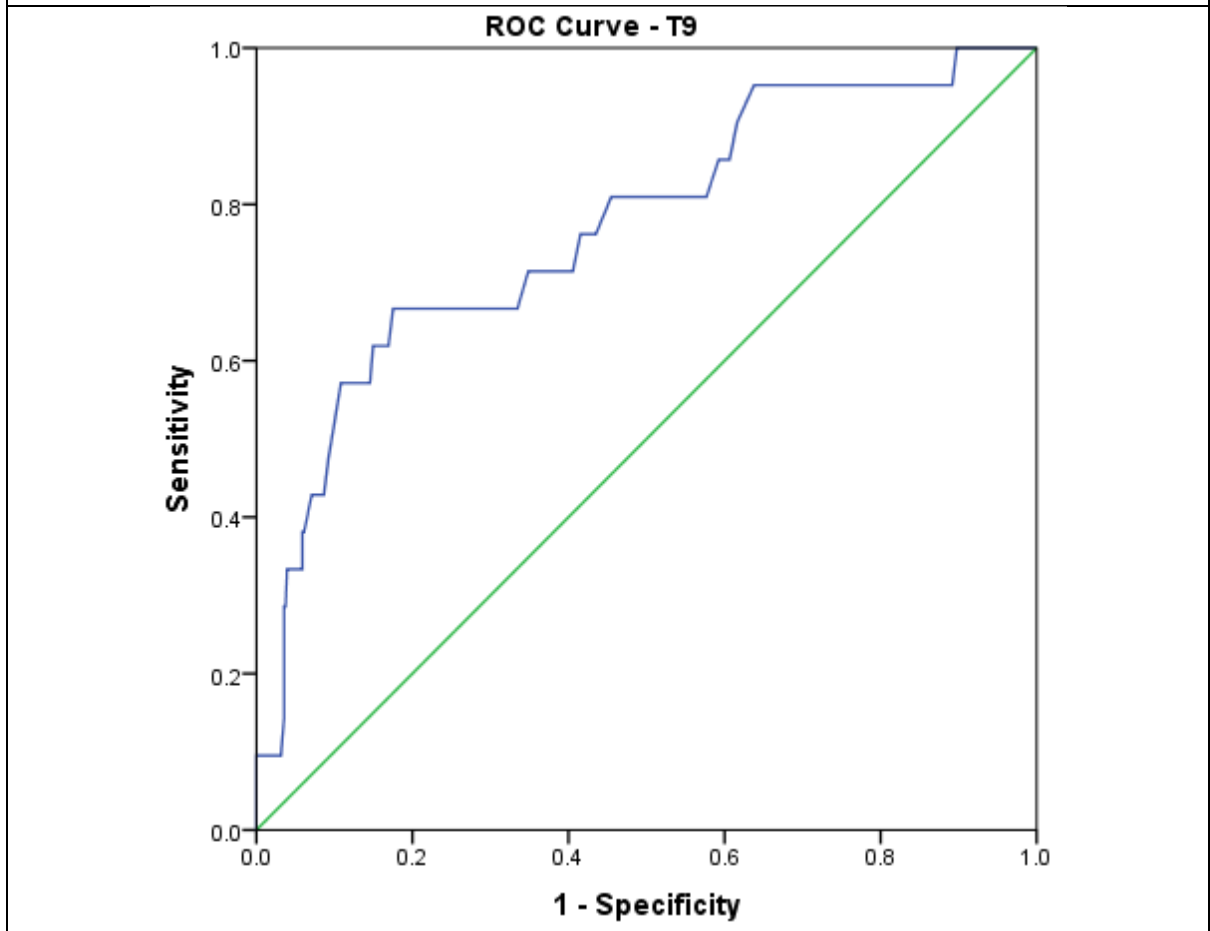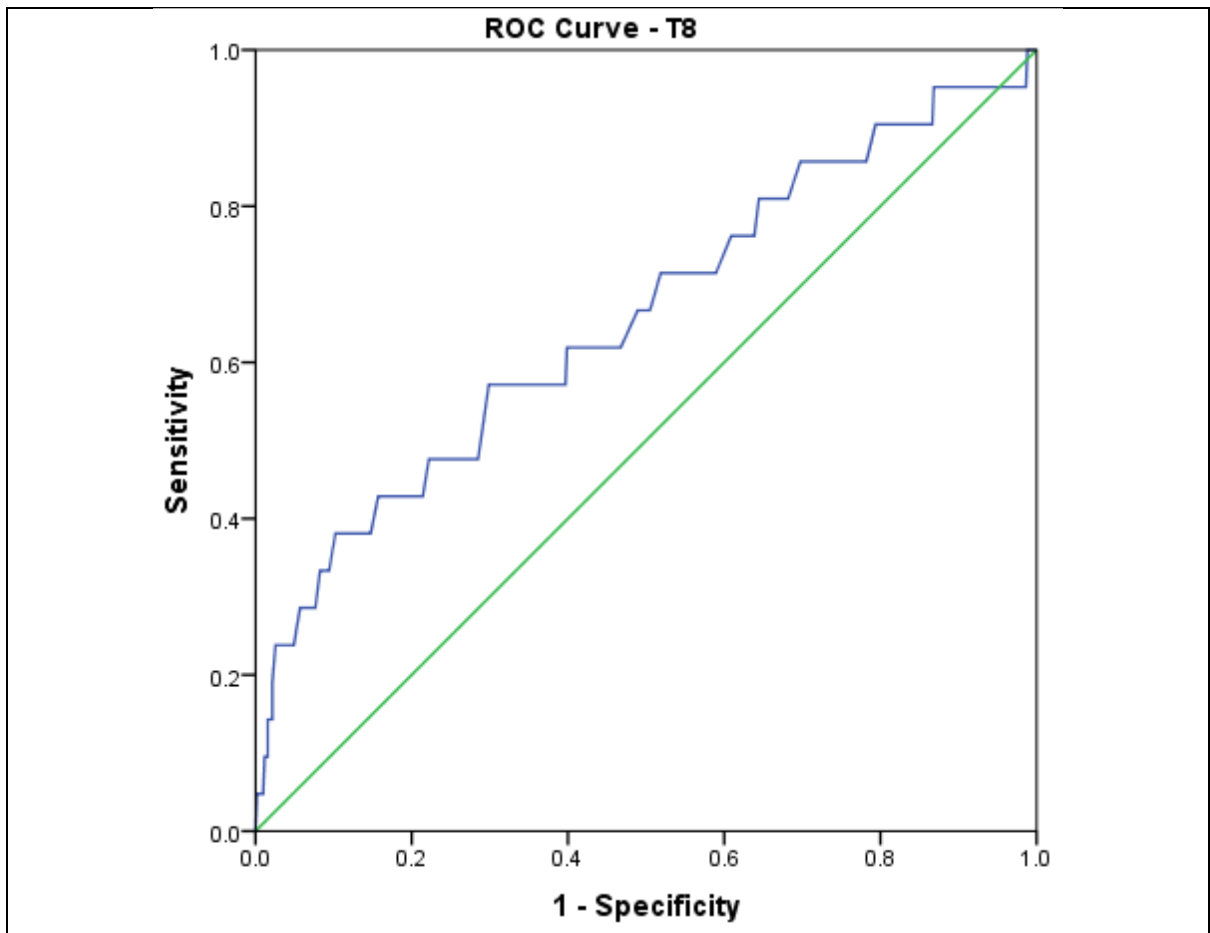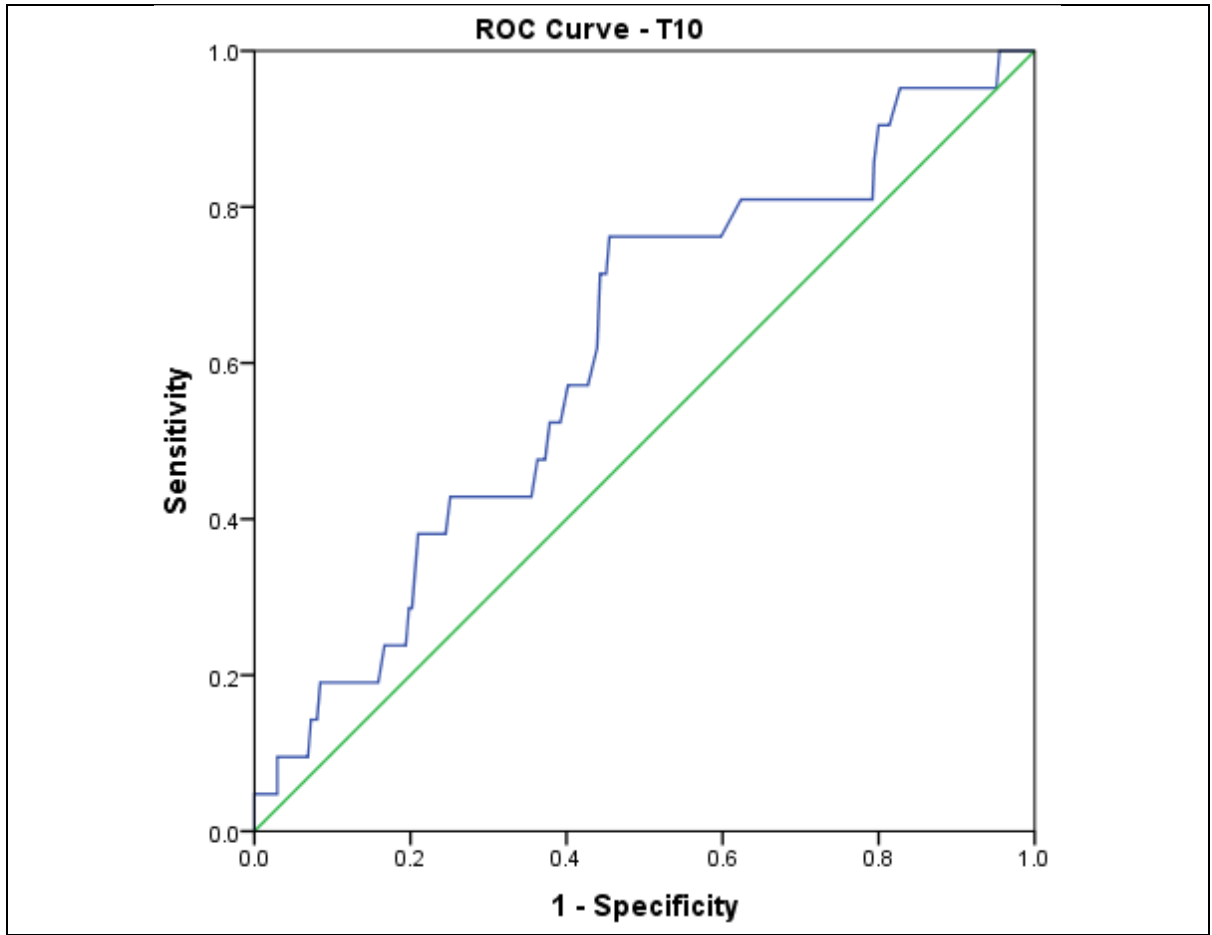
ROC Curve- T2

ROC Curve - T3

ROC Curve - T4

ROC Curve - T5

ROC Curve - T6

ROC Curve - T7

ROC Curve - T8

ROC Curve - T9

ROC Curve - T10

ROC Curve - T11

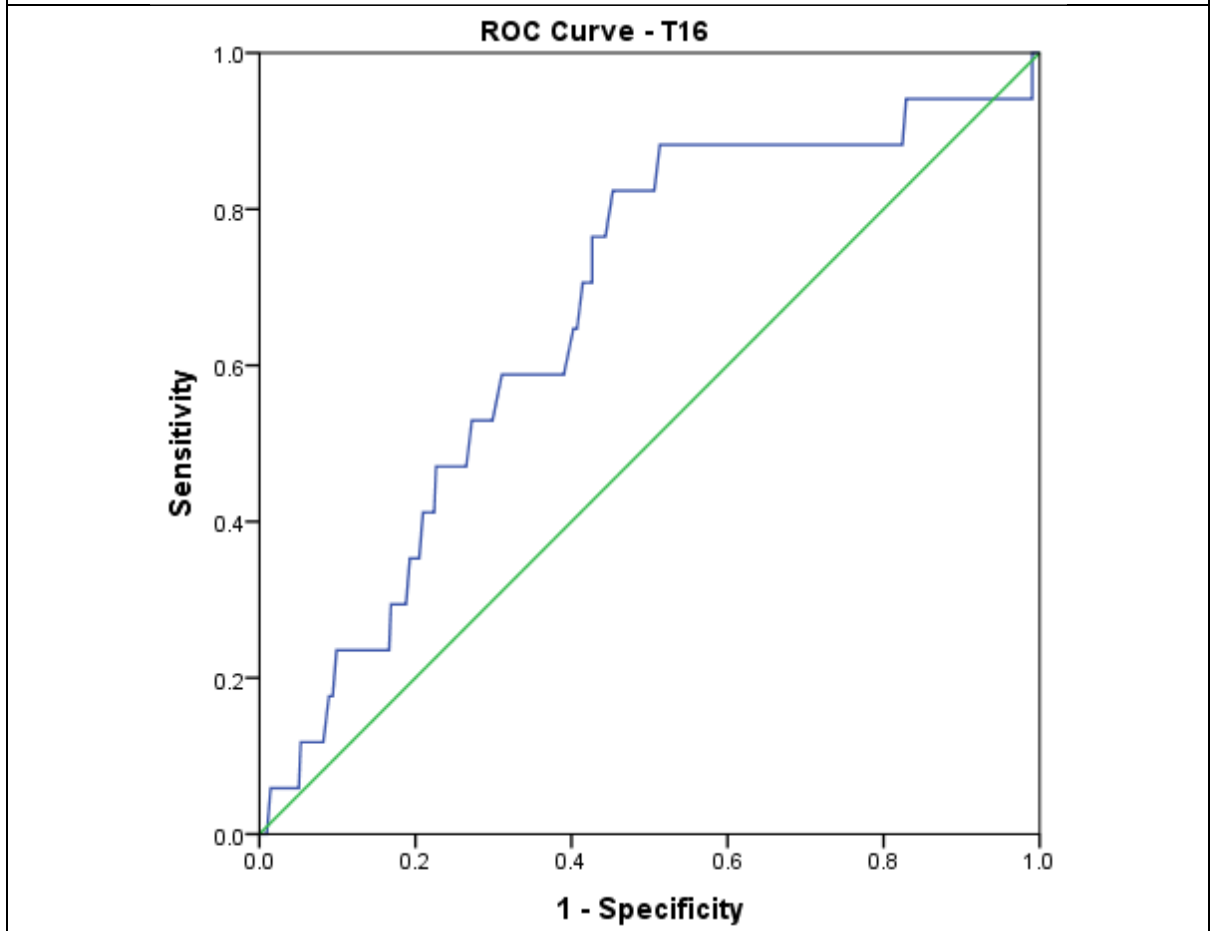ROC Curve - T12

ROC Curve - T13

ROC Curve - T14
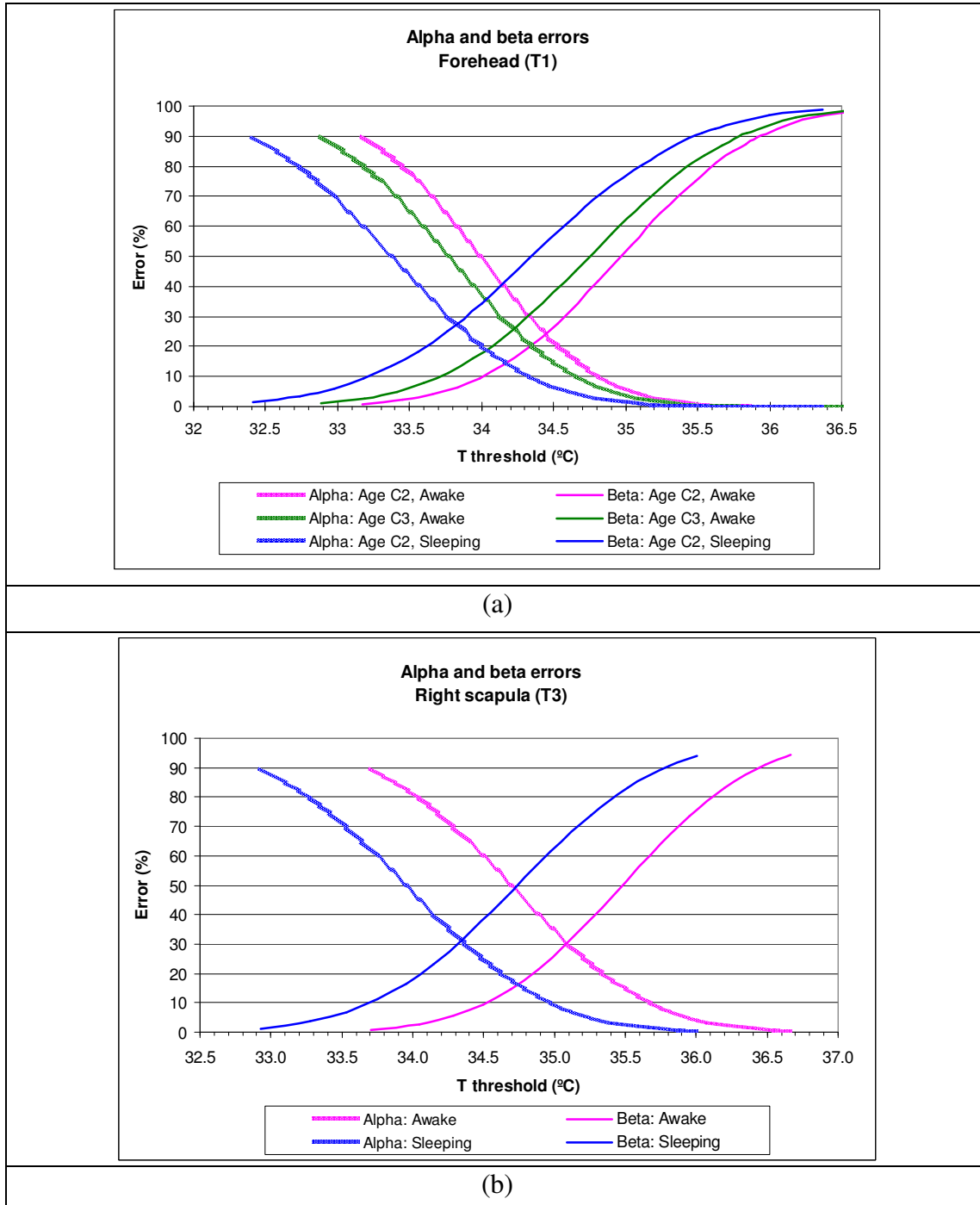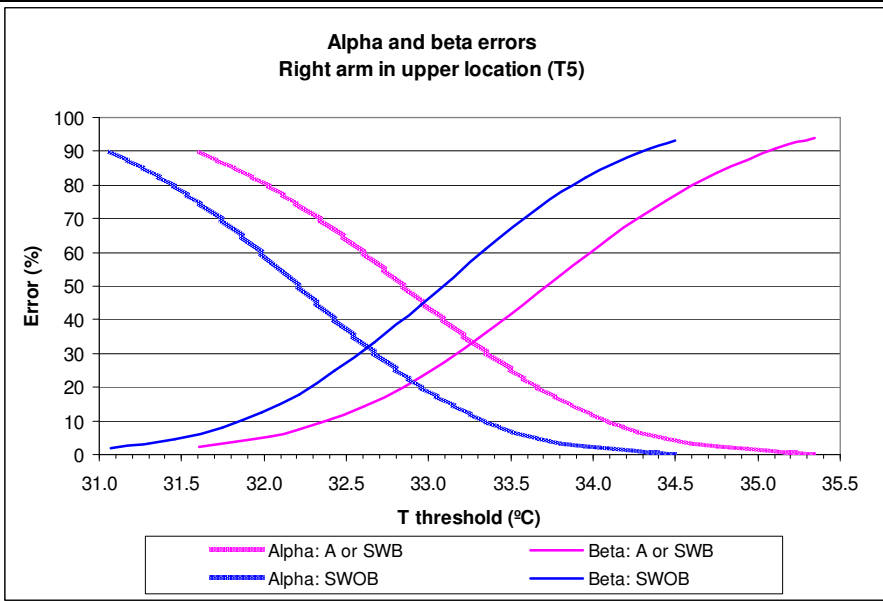
ROC Curve - T15
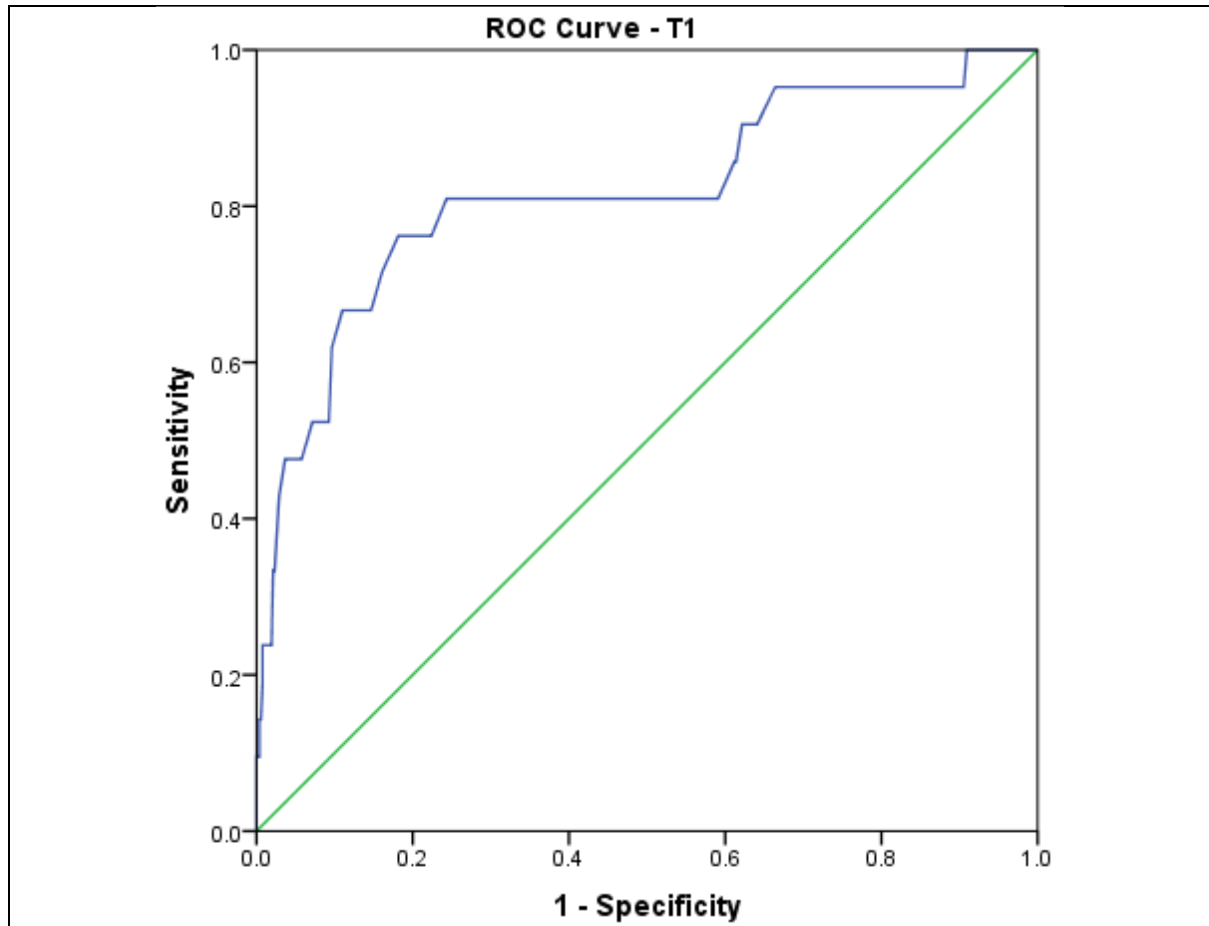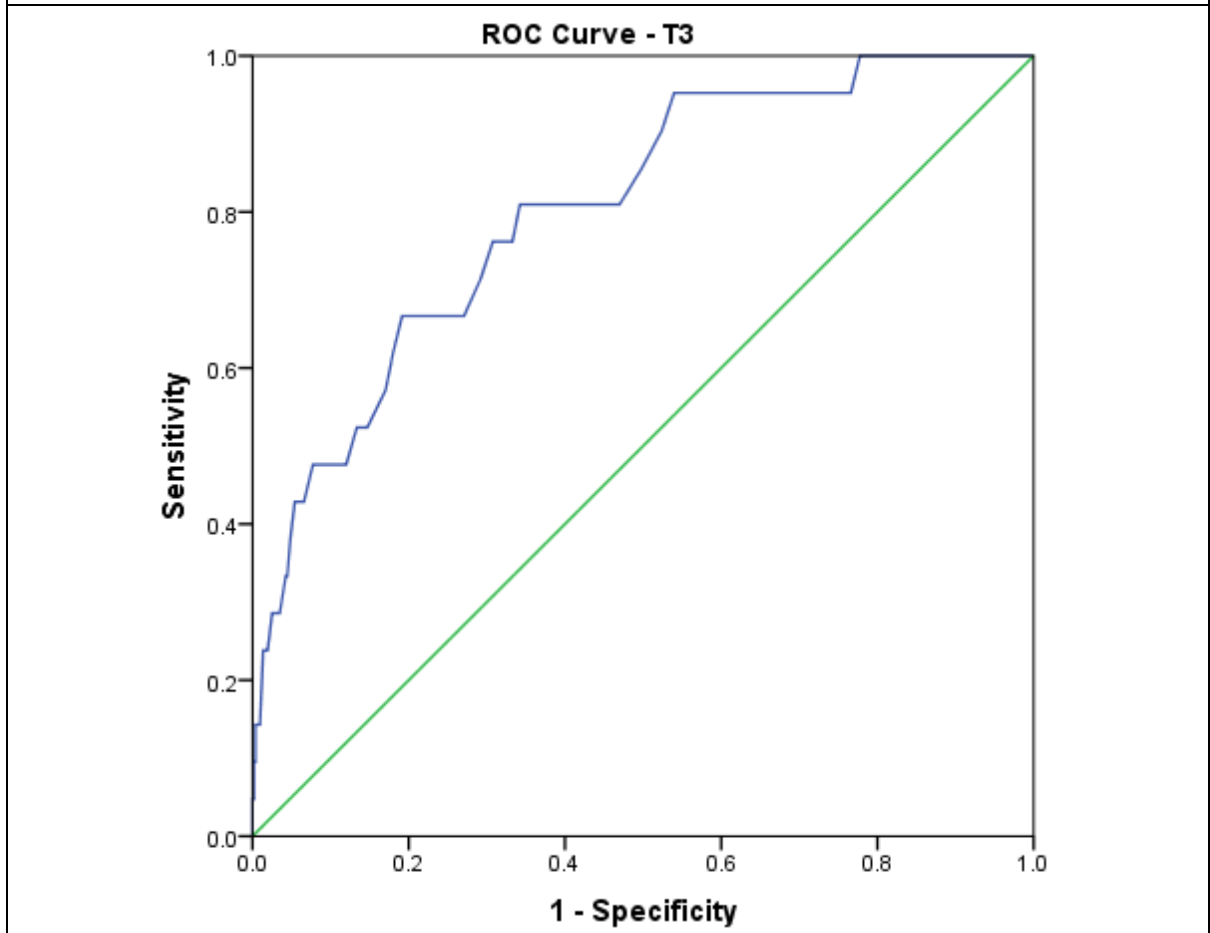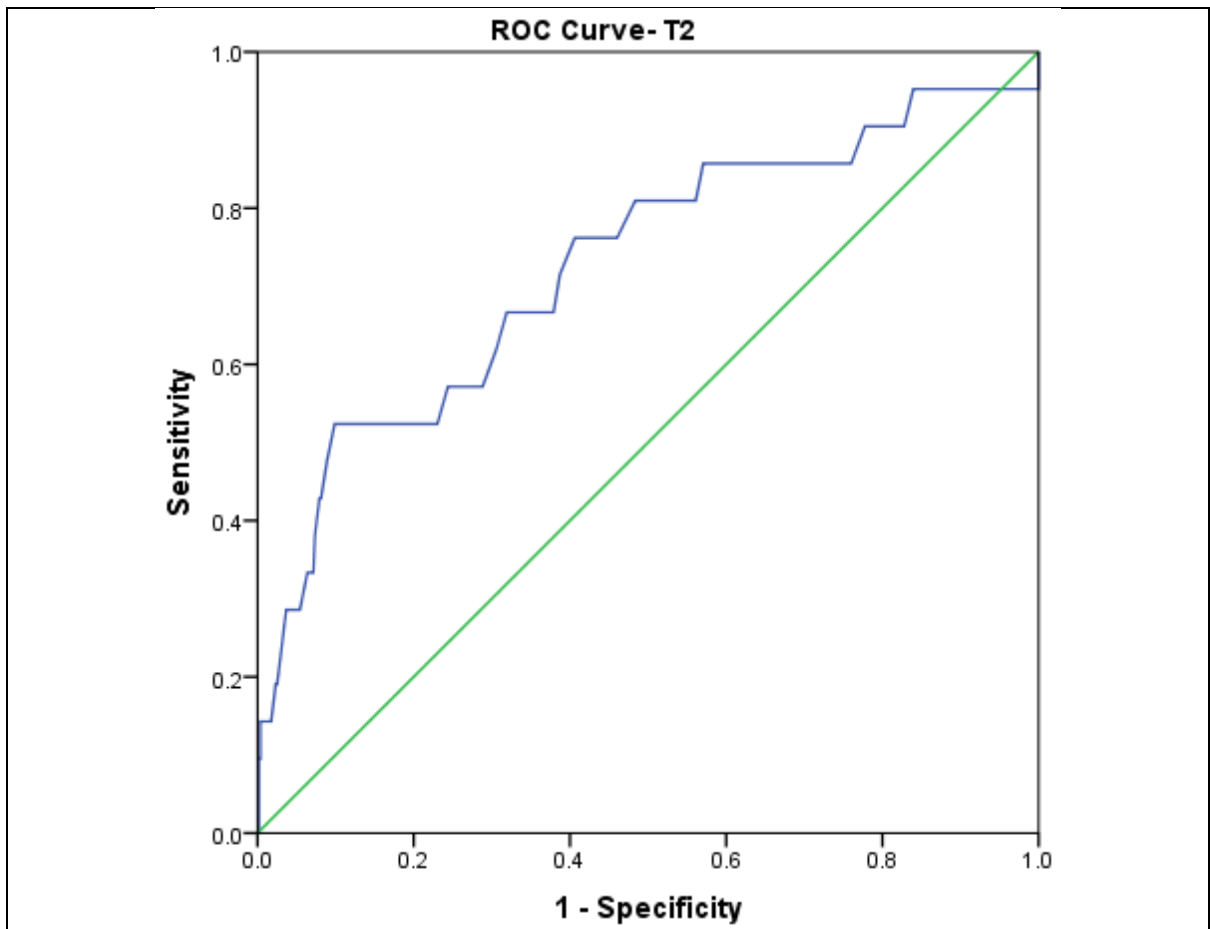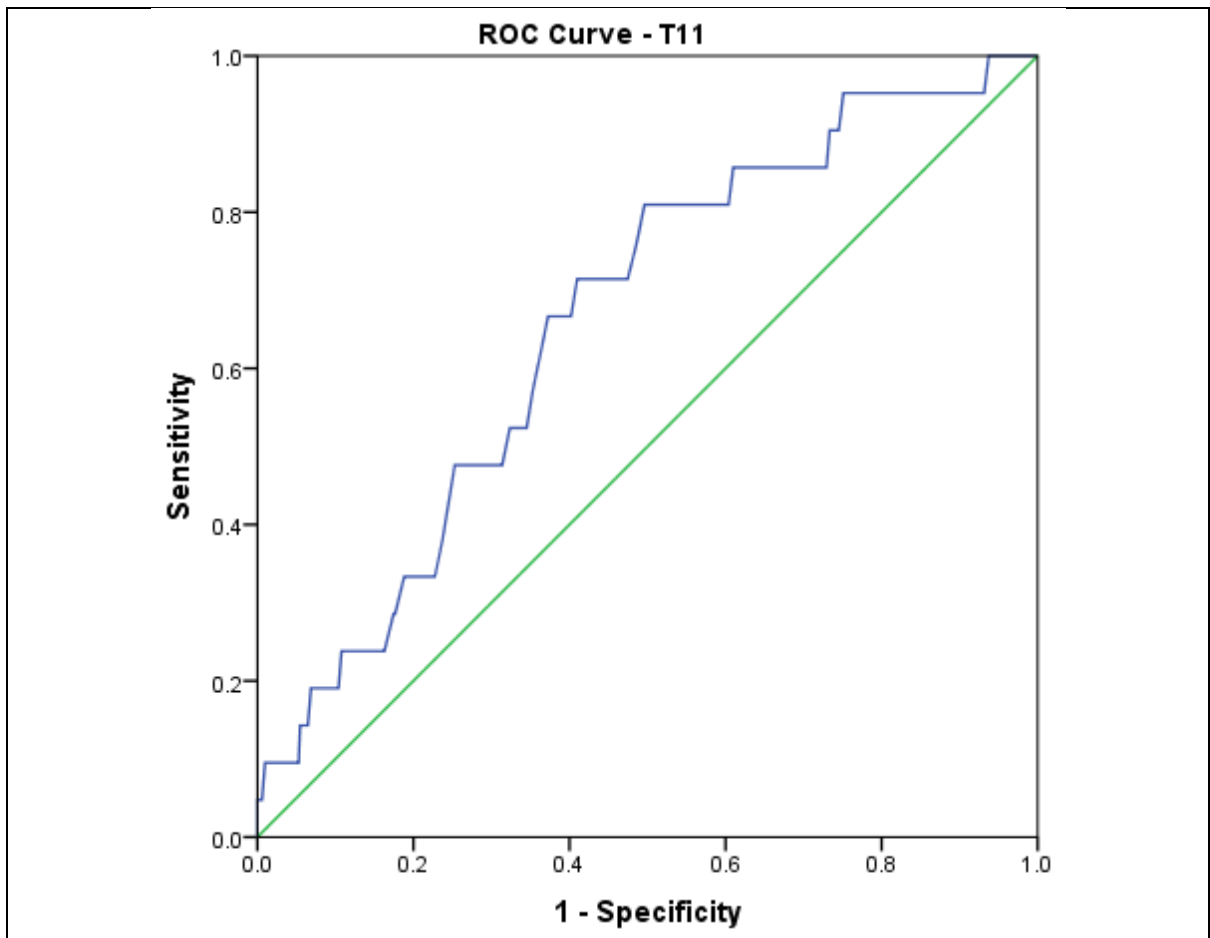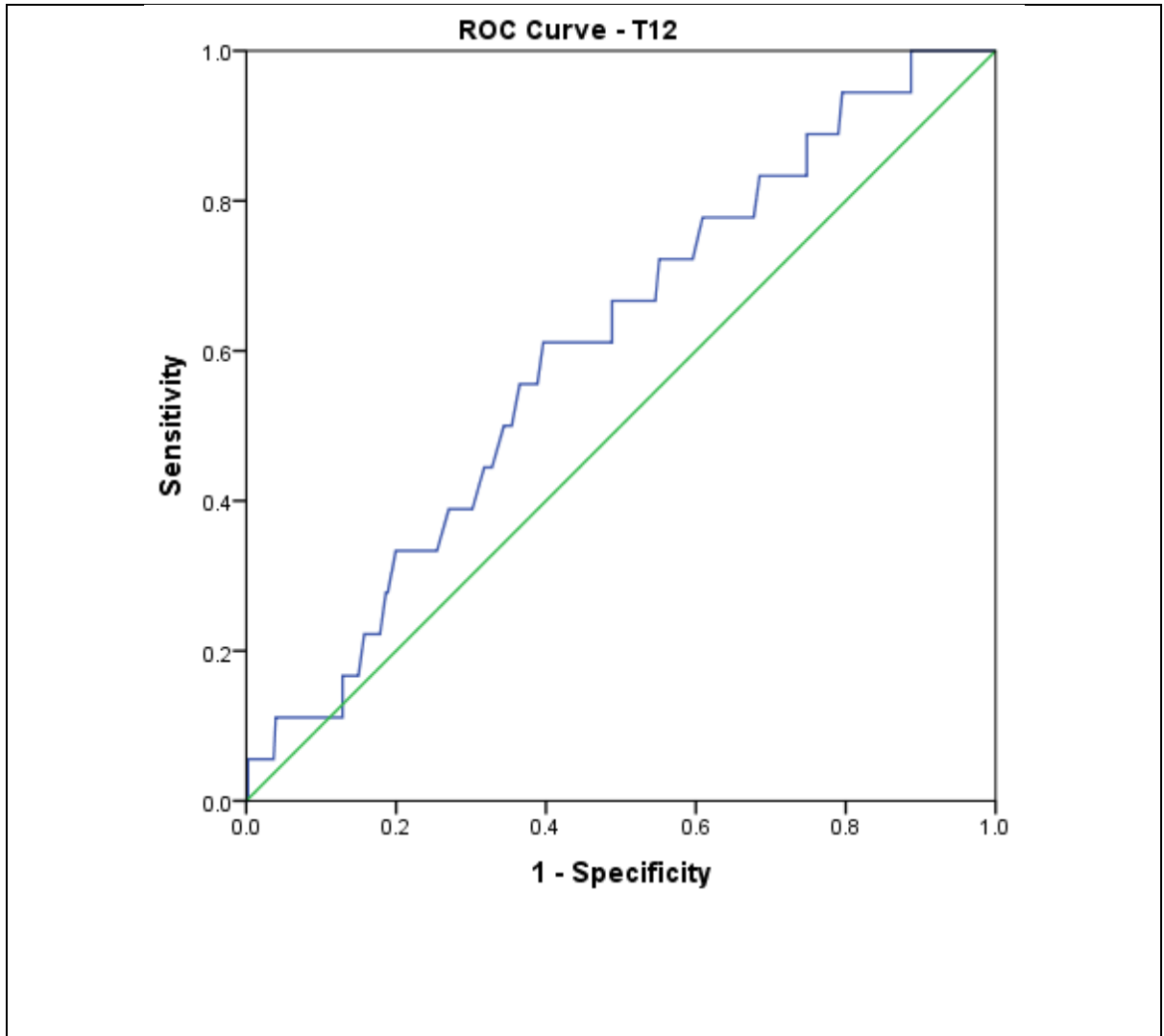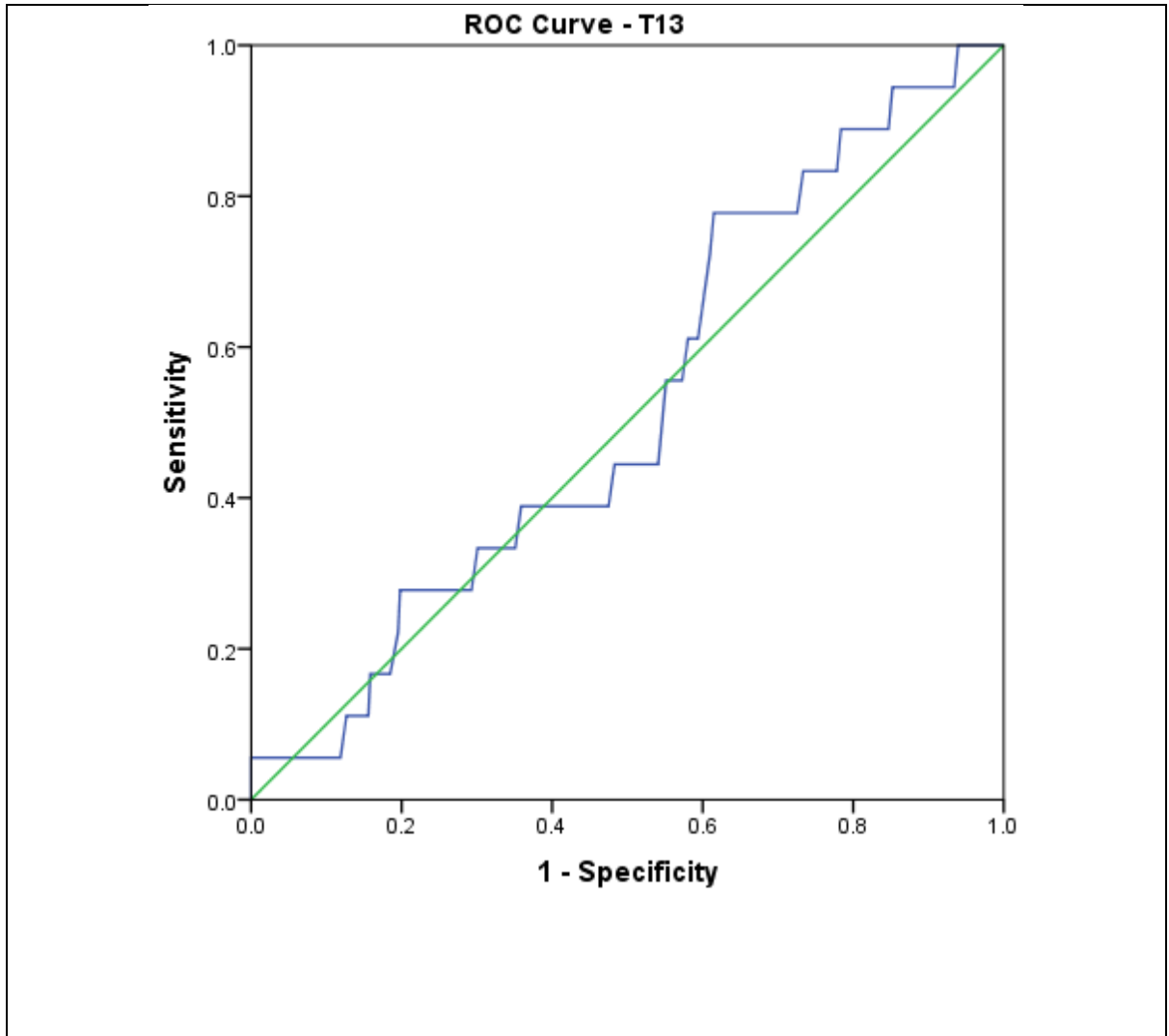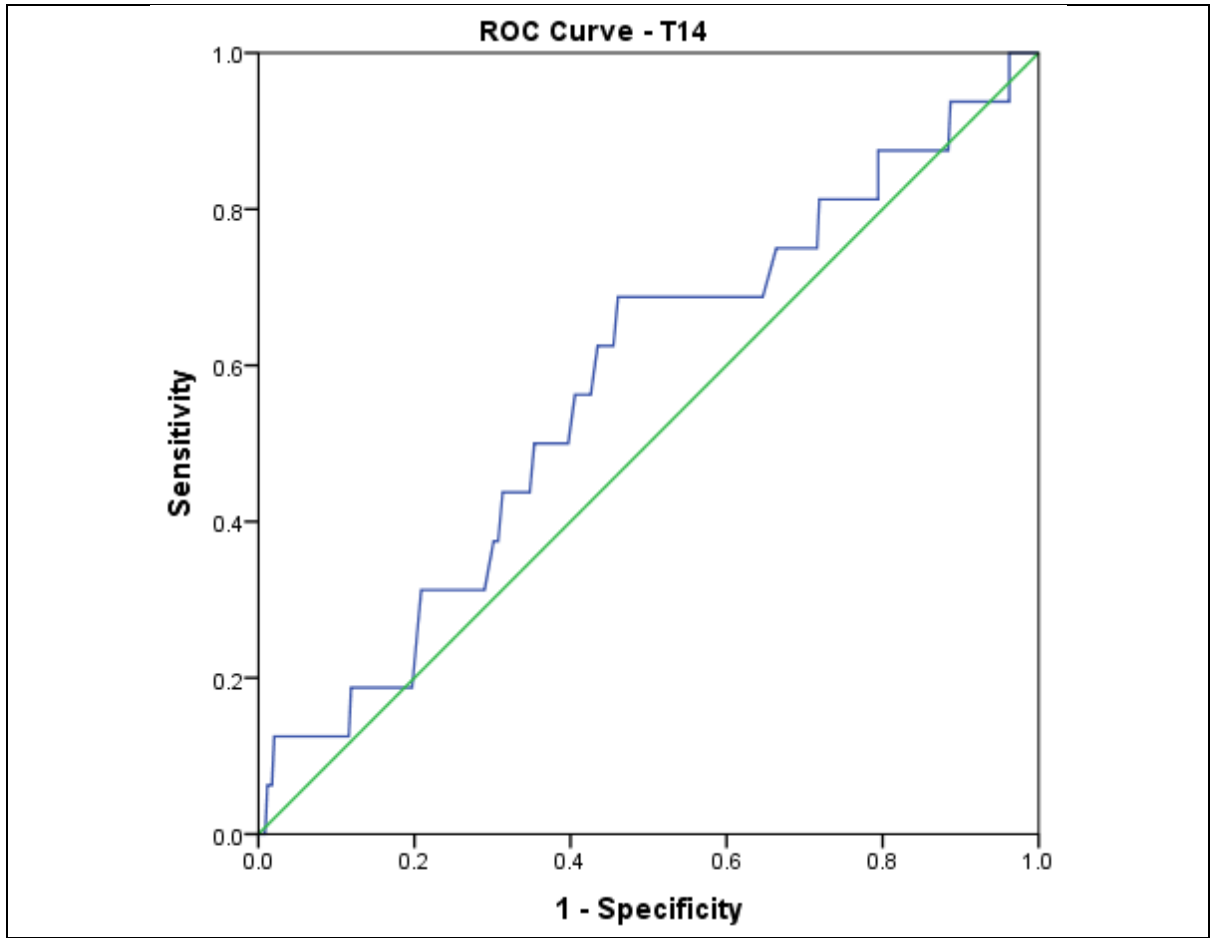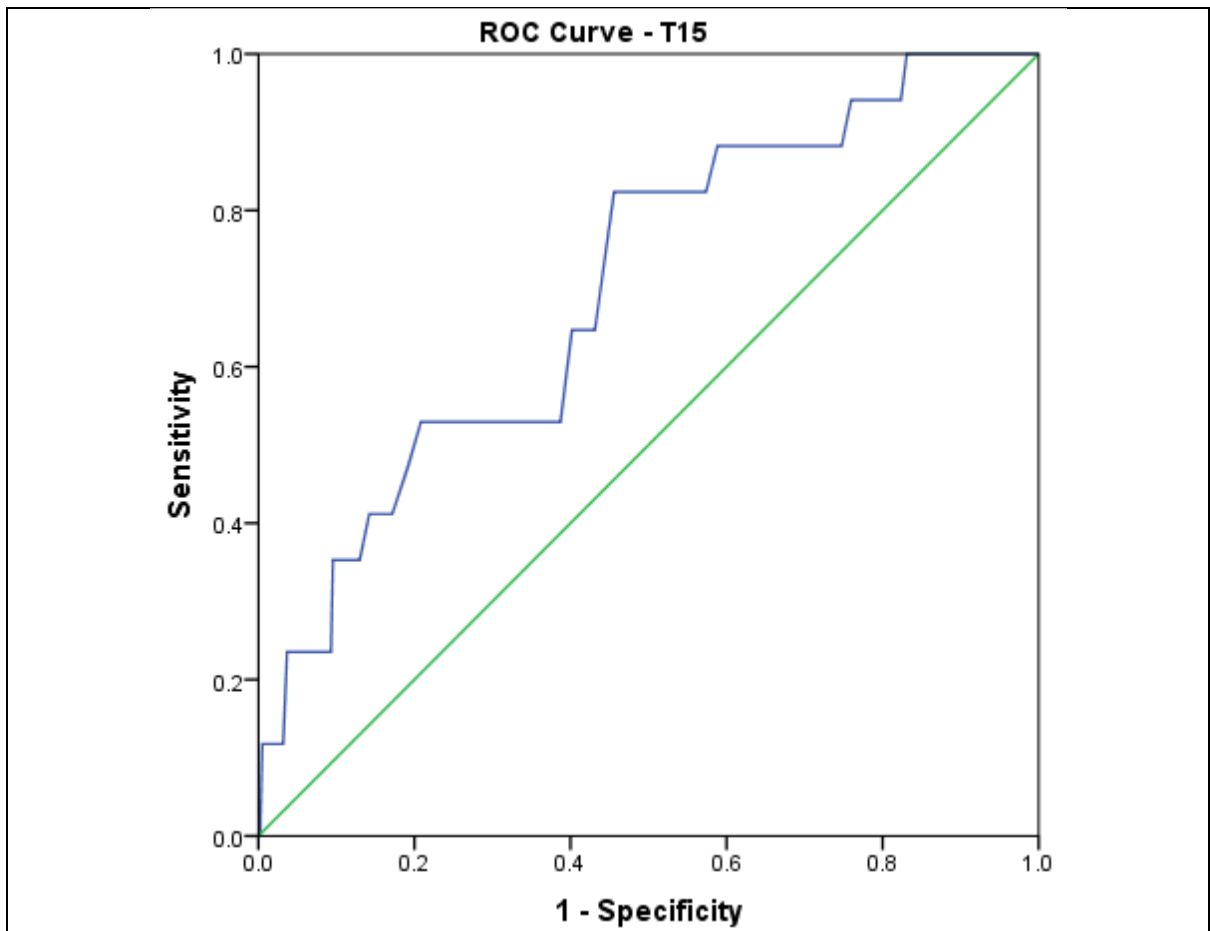
ROC Curve - T16

Figure A1. ROC curves for all the studied local skin temperatures as individual markers of fever.
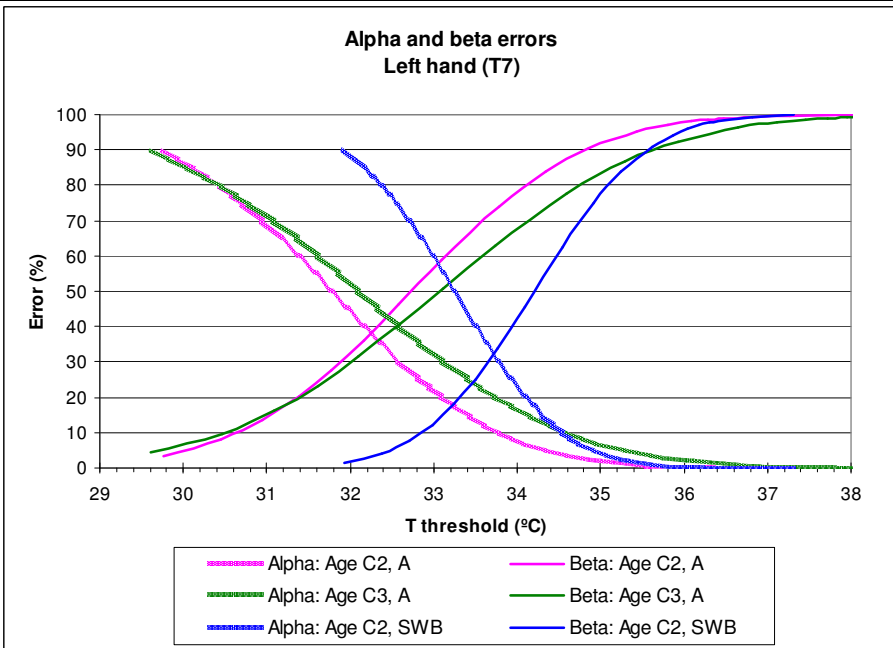
Appendix B

Alpha and beta curves obtained for potentially suitable locations for fever monitoring and detection (**Figure B1**).
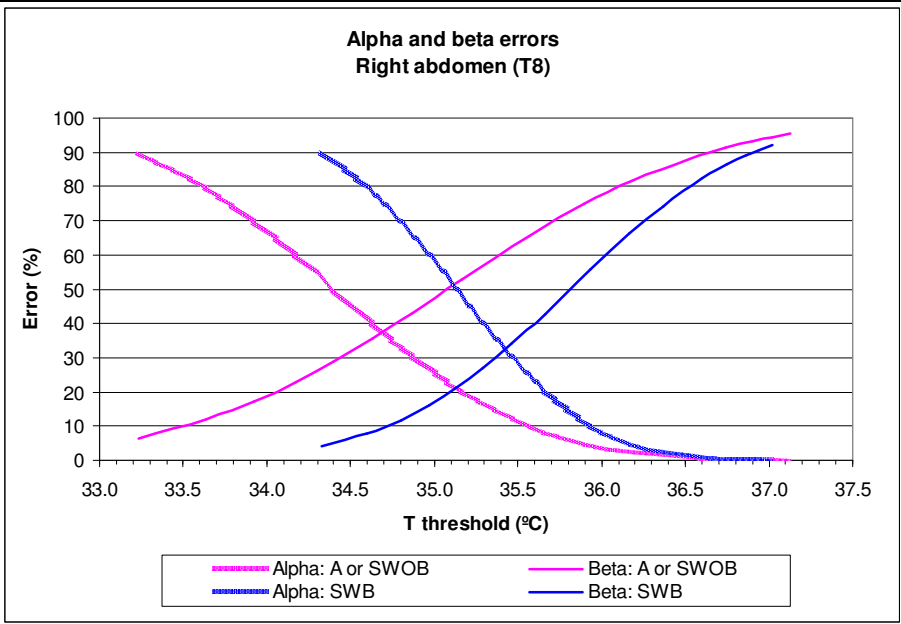

(a)

(b)

Alpha and beta errors
Right arm in upper location (T5)

Legend:
- Alpha: A or SWB
- Beta: A or SWB
- Alpha: SWOB
- Beta: SWOB

(c)

Alpha and beta errors
Left hand (T7)

Legend:
- Alpha: Age C2, A
- Beta: Age C2, A
- Alpha: Age C3, A
- Beta: Age C3, A
- Alpha: Age C2, SWB
- Beta: Age C2, SWB

(d)

**Alpha and beta errors
Right abdomen (T8)**
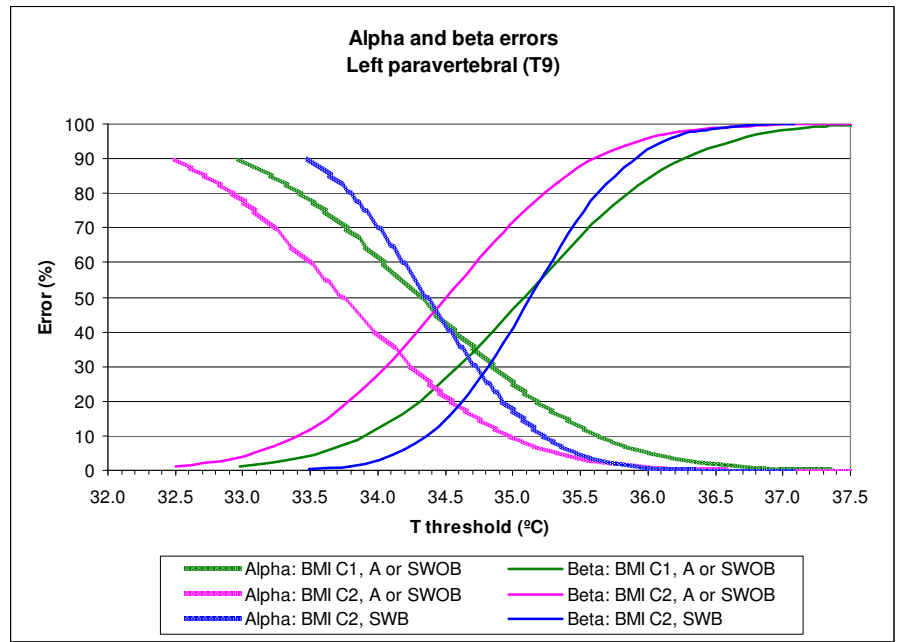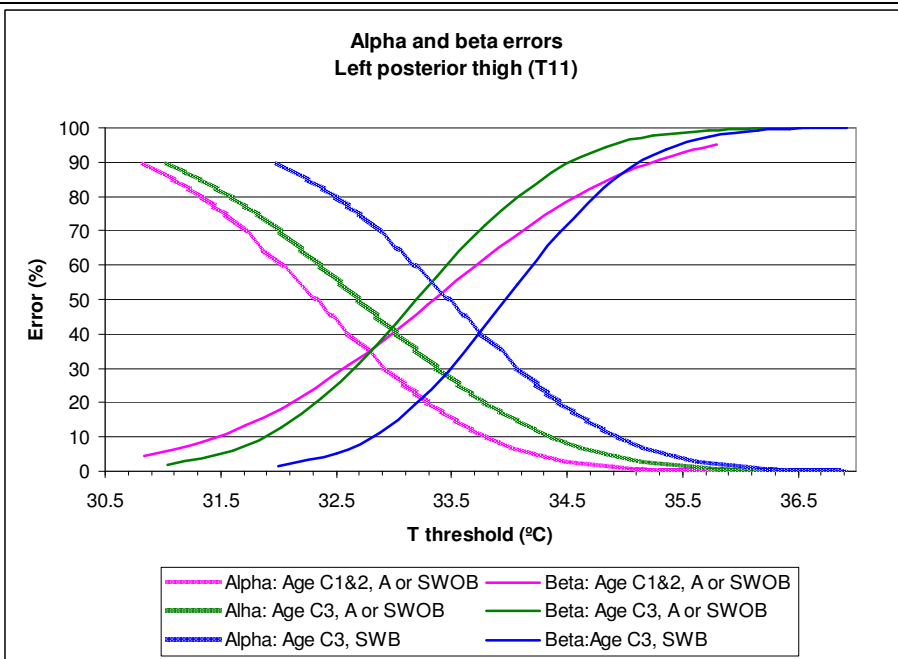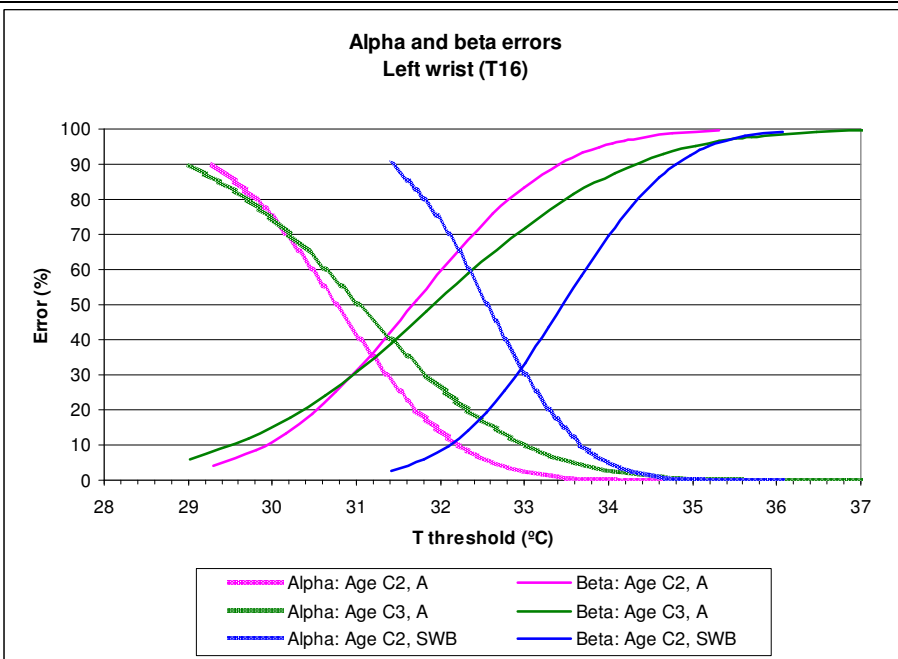
(e)

**Alpha and beta errors
Left paravertebral (T9)**

(f)

(g)

(h)

Figure B1: Alpha (detected false positives) and beta (missed true positives) errors for different skin temperature thresholds at (a) forehead for 3/6 relevant sub-cohorts; (b) right scapula; (c) right arm in upper location; (d) left hand for 3/9 relevant sub-cohorts; (e) right abdomen; (f) left paravertebral for 3/6 relevant sub-cohorts; (g) left posterior thigh for 3/4 relevant sub-cohorts; and (h) left wrist for 3/9 relevant sub-cohorts.

Nomenclature: (*A)* awake; (*SWB*) sleeping with a blanket; (*SWOB*) sleeping without a blanket; (Age C1) 0 – 8 months; (Age C2) 8 – 16 months; (Age C3) 8 – 16 months; (BMI C1) Underweight; (BMI C2) Normal healthy.

References

[1] National Institute for Health and Care Excellence – NICE (22nd May2013) – *NICE Clinical Guideline 160 Fever in under 5s: assessment and initial management.* http://guidance.nice.org.uk/CG160

[2] Kocoglu H, Goksu S, Isik M, Akturk Z and Bayazit WA (2002) *Infrared tympanic thermometer can accurately measure the body temperature in children in an emergency room setting.* Int J Pediatric Otorhinolaryngology 65: 39–43

[3] Cascetta F (1995) *An evaluation of the performance of an infrared tympanic thermometer.* Measurement 16(4): 239-246

[4] Nimah MM, Bshesh K, Callahan JD and Jacobs BR (2006) *Infrared tympanic thermometry in comparison with other temperature measurement techniques in febrile children.* Pediatric critical care medicine 7(1): 48-55

[5] van Staaija BK, Roversa MM, Schilderb AG and Hoesa AW (2003) *Accuracy and feasibility of daily infrared tympanic membrane temperature measurements in the identification of fever in children.* Int J Pediatric Otorhinolaryngology 67: 1091-1097

[6] Lamonaca F, Polimeni G, Barbé K, Grimaldi D (2015) *Health parameters monitoring by smartphone for quality of life improvement.* Measurement 73: 82-94

[7] Atallah L, Lo B, Ali R, King R, Yang GZ (2009) *Real-Time Activity Classification Using Ambient and Wearable Sensors.* IEEE Transactions on Information Technology In Biomedicine 13(6): 1031-1039

[8] Mansor H, Shukor MHA, Meskam SS, Rusli NQAM and Zamery NS (2013) *Body temperature measurement for remote health monitoring system.* 2013 *IEEE* International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA)*,* Kuala Lumpur, pp: 1-5

[9] Ray PP (2016) *Internet of Things cloud enabled MISSENARD index measurement for indoor occupants.* Measurement 92: 157-165

[10] Ogai K, Fukuoka M, Kitamura KI, Uchide K, Nemoto T (2016) *Development of a small wireless device for perspiration monitoring.* Medical Engineering & Physics 38(4): 391-397

[11] Milici S, Amendola S, Bianco A and Marrocco G (2014) *Epidermal RFID passive sensor for body temperature measurements.* 2014 IEEE RFID Technology and Applications Conference (RFID-TA), Tampere, pp: 140-144

[12] Affanni A, Chiorboli G (2015) *Design and characterization of a real-time, wearable, endosomatic electrodermal system.* Measurement 75: 111-121

[13] McKenzie M, Gainor M. *Releasable clothing with temperature sensor for bedridden patients*. U.S. Patent 5 802 611 A, Sep 8, 1998.

[14] Kaplan RB, Johnson TM, Schneider RO, Krishnan SM. (2011) *A design for lowcost and scalable non-contact fever screening system.* ASEE Annual Conference and Exposition, June 26-29, Vancouver, Canada.

[15] Rubinstein E. *Fever alarm system.* U.S. Patent 6 547 745 B1, April 15, 2003.

[16] Garcia-Souto MDP, Dabnichki P (2013) *Skin temperature distribution and thermoregulatory response during prolonged seating.* Building and Environment 69: 14-21

[17] Garcia-Souto MDP, Dabnichki P (2016) *Core and local skin temperature: 3-24 months old toddlers and comparison to adults.* Building and Environment 104: 286-295

[18] Kimberger O, Thell R, Schuh M, Koch J, Sessler DI, Kurz A (2009) *Accuracy and precision of a novel non-invasive core thermometer.* Br J Anaesth 103 (2): 226-231

[19] Kitamura KI, Zhu X, Chen WX, Nemoto T (2010) *Development of a new method for the noninvasive measurement of deep body temperature without a heater.* Medical Engineering & Physics 32(1): 1-6

[20] Ring EFJ, Jung A, Zuber J, Rutowski P, Kalicki B, Bajwa U (2008) *Detecting Fever in Polish Children by Infrared Thermography.* 9th Int Conf on Quantitative Infrared Thermography, July 2-5, Krakow, Poland.

[21] Bridges E and Thomas K (2009) *Noninvasive Measurement of Body Temperature in Critically Ill Patients.* Critical Care Nurse 29(3): 94-97

[22] Langham GE, Maheshwari A, Contrera K, You J, Mascha E and Sessler DI (2009) *Noninvasive Temperature Monitoring in Postanesthesia Care Units.* Anesthesiology 111(1): 90-96

[23] Lawson L, Bridges EJ, Ballou I, Eraker R, Greco S, Shively J, Sochulak V (2007) *Accuracy and precision of noninvasive temperature measurement in adult intensive care patients.* American Journal of Critical Care 16(5): 485-496

[24] Terndrup TE, Milewski A (1991) *The performance of 2 tympanic thermometers in a pediatric emergency department.* Clinical Pediatrics 30(4): 18-23 Suppl. S. Conference Information: Conference on Fever and Tympanic Thermometry, Jan 18, 1991 San Diego, CA

[25] Kelly B and Alexander D (1991) *Effect of Otitis Media on Infrared Tympanic Thermometry.* Clinical Pediatrics 30(4): 46-48 Suppl. S. Conference Information: Conference on Fever and Tympanic Thermometry, Jan 18, 1991 San Diego, CA

[26] Shenep JL, Adair JR, Hughes WT, Roberson PK, Flynn PM, Brodkey TO, Fullen GH, Kennedy WT, Oakes LL and Marina NM (1991) *Infrared, thermistor, and glass-mercury thermometry for measurement of body-temperature in children with cancer.*

Clinical Pediatrics 30(4): 36-41 Supplement: Suppl. S Conference Information: Conference on Fever and Tympanic Thermometry, Jan 18, 1991 San Diego, CA

[27] International Standard Organization. *ISO 9886:2004(E), Ergonomics – Evaluation of thermal strain by physiological measurements.*

[28] Nagano K, Takaki A, Hirakawa M and Tochihara Y (2005) *Effects of ambient temperature steps on thermal comfort requirements*. Int J Biometeorololy 50: 33-39

[29] Metz CE (1978) *Basic principles of ROC analysis.* Sem Nuc Med. 8:283-298

[30] Hogan DH, Shipman S, Smith K (2015) *Simple infrared thermometry in fever detection: Consideration in mass fever screening.* Am J Disaster Med 10(1):69-74

[31] Deng ZS and Liu J (2008) *Effect of Fixing Material on Skin-Contact Temperature Measurement by Wearable Sensor* IEEE 5th International Summer School and Symposium on Medical Devices and Biosensors: 151-154