Infrared Thermal Imaging to Detect Inflammatory Intra-Abdominal Pathology in Infants

[1] CONSTANCE BARSON, [2*] REZA SAATCHI, [2,3] PRASAD GODBOLE,
[3] SHAMMI RAMLAKHAN
*Corresponding author: r.saatchi@shu.ac.uk

[1] UNIVERSITY OF SHEFFIELD, SHEFFIELD, S10 2TN, UNITED KINGDOM.
[2] SHEFFIELD HALLAM UNIVERSITY, SHEFFIELD, S1 1WB, UNITED KINGDOM.
[3] SHEFFIELD CHILDREN'S HOSPITAL, SHEFFIELD, S10 2TH, UNITED KINGDOM.

Abstract: - A thermal imaging method to detect inflammatory intra-abdominal pathology in infants is proposed and evaluated through a clinical trial. Nine surgical infants, mean chronological age 58 days old (range: 21-83 days), mean weight 2.65 kg (range: 2.45-3.15 kg) with abdominal pathologies were included in the analysis. Infrared thermal image processing consisted of selecting the surgical region of interest where the area of abdominal inflammation was most likely to be, and an abdominal reference region on the same infant, with the aid of clustering segmentation. Skewness was found to be the most sensitive variable to significantly differentiate between the surgical region and reference region (p = 0.022). Multilinear regression analysis indicated that the relationship between the temperature difference signified by skewness and the patients' demographic information (age at time of imaging, gestational age at birth, weight at the time of imaging, birthweight, last stool prior to imaging and last oral intake prior to imaging) was not significant. The study indicated that inflammatory regions, such as those found in infants following surgery, would have a significantly different temperature distribution than the surrounding skin. The method differentiated between an inflammatory and non-inflammatory region on the abdomen.

1 Introduction

Infrared thermal imaging (IRT) in the medical field is a rapidly advancing technology and its possible clinical applications are continually expanding [1]. IRT imaging could be a possible complementary technology to conventional radiography in that it is a contactless and harmless method that is capable of sensitive, reliable and accurate detection of temperature changes indicative of disease [2].

The acute inflammatory response which occurs with infection is accompanied by perfusion alterations, such a temperature increase or decrease, may be detected as a 'hot spot' or 'cold spot' on a thermogram respectively [3]. The thermal map created could indicate the onset of infection before clinical signs are evident and may allow for the development of indicators for surgical intervention, further reinforcing the potential of IRT imaging as an early diagnostic tool.

IRT imaging has been used to measure respiration rate in a non-contact manner, i.e. without any sensor attached to the subject's body [4], [5]. It has proved valuable for detecting diseases with an infectious or inflammatory aetiology, or where a clinical abnormality may cause blood flow to be reduced or increased [6].

In pediatric, IRT has allowed detection of vertebral fracture in osteogenesis imperfecta [7], assisted with diagnosis of limp [8], differentiated between wrist fracture and wrist sprain [9] and provided with an accurate objective measure of temperature changes associated with juvenile idiopathic arthritis [10]. There were also studies demonstrating the potential of IRT in infants for monitoring and diagnosis [1], [11]-[16].

Few research studies have been conducted in the field of IRT imaging on premature infants housed in convection incubators [17]. This has been a challenging pursuit for a multitude of reasons, particularly as premature infants have a high level of heat loss due to immature thermoregulation [16]. Preterm infants are at a higher risk of heat loss due to a disproportionate body mass-to-surface area ratio, inadequate amounts of brown adipose tissue and proteins to enable metabolic heat production, relatively thin skin and increased permeability,

Constance Barson,
Reza Saatchi, Prasad Godbole,
Shammi Ramlakhan

leading to high levels of evaporative loss and immature vasomotor control and homeostatic regulation [18].

As the skin of neonates is much thinner than the skin in older children and neonates have little abdominal subcutaneous fat, it is postulated that generating an accurate thermal map of the neonatal abdomen using IRT may be of clinical use for the detection of intra-abdominal infection in neonates [16], [19].

The aim of this study was to determine whether dynamic passive IRT imaging could assist clinicians in detecting inflammatory intra-abdominal pathology. This could have applications for monitoring and detection of the conditions such as such as Necrotizing Enterocolitis (NEC) in neonates (NEC).

NEC is one of the most severe inflammatory conditions of the gastrointestinal tract in neonates. This complex and multifactorial condition is predominantly a disorder of the premature infant [20], and usually presents within the first 2-3 weeks of life. The ability to be forewarned of impending early stage NEC so far evades clinicians [21]. Surgical and medical neonates are prone to contract infections which can be difficult to detect in their early stages and such infections are more likely to cause septicaemia in premature neonates than in the older age groups [21], [22]. Early intervention in those neonates with fulminant or severe NEC is paramount and currently abdominal radiographs and ultrasonography have limited diagnostic impact in early stage or uncomplicated NEC [23], [24]. A feasibility study was conducted to assess abdominal skin temperature using IRT imaging in extremely low birth weight (ELBW) infants to determine whether a relationship existed between abdominal skin temperature and NEC [14]. Concurrent thermal images and abdominal radiographs were obtained in 10 infants and only three had radiographic evidence of NEC. Infants with radiographic NEC had lower abdominal skin temperature (35.3 ± 0.8 °C) compared to those without evidence of NEC (36.6± 0.9 °C), which reflects the picture of ischaemia and inadequate perfusion of the bowel seen in NEC.

In the following sections the study's methodology and results are described.

2 Methodology

2.1 Ethical Approval

Prior to commencing the study, ethical approvals were received from the National Health Service (NHS) Research Ethics Committee, United Kingdom and the Research and Development Department at the University of Sheffield, United Kingdom. The study was conducted in compliance with the protocol, Good Clinical Practice (GCP) [25] and research ethics.

2.2 Inclusion and Exclusion Criteria for Participants

Inclusion criteria:
- Admitted to the Neonatal Surgical Unit at 37 weeks gestation or older.
- Received surgical intervention for a gastrointestinal condition involving the gastrointestinal tract from the pylorus to the anus.

Exclusion criteria:
- Infants who were unstable or very unwell were excluded to avoid any further stress to them and their parents.
- Infants with any abdominal covering which occluded the area to be imaged which for various reasons could not be removed.
- Any major congenital intestinal abnormality, congenital heart disease or conditions which might have altered thermoregulation in the infants. Congenital heart disease, especially cyanotic anomalies, may reduce bowel perfusion and cause intestinal ischaemia [26]. Common conditions which alter thermoregulation include infection, sepsis, hypoglycaemia, metabolic impairment, abnormal skin integrity, neural tube defects, neurological impairment and growth restriction [27].
- Inability of carers (parents etc.) to communicate in English. The patient information sheet and consent form were only available in English.
- Inability of carers to give informed consent.

2.3 Recruitment

The carers of all infants included in the study gave informed consent for participation and sequences of thermal images were obtained from all of these participating infants. Enrolment resulted in a cohort of 10 abdominal surgical infants, however infant number 9 was excluded from the analysis at later stage as the abdomen region was partially occluded. For the 9 infants included in the analysis, their demographic information and surgical conditions are presented in Appendix A.

Constance Barson,
Reza Saatchi, Prasad Godbole,
Shammi Ramlakhan

2.4 Data Recording

The thermal camera used in this study was the FLIR T630sc handheld camera. The camera's temperature sensitivity is to within 40 mK, image resolution 640×480 pixels, spectral range 7.5 μm to 13 μm, dynamic range 14 bits and operating temperature range -14 °C to 50 °C. It was connected via a USB link to a high specification laptop to store data. The thermal videos were processed and analyzed in Matlab© [28]. Figure 1 shows a schematic of the data recording set up.
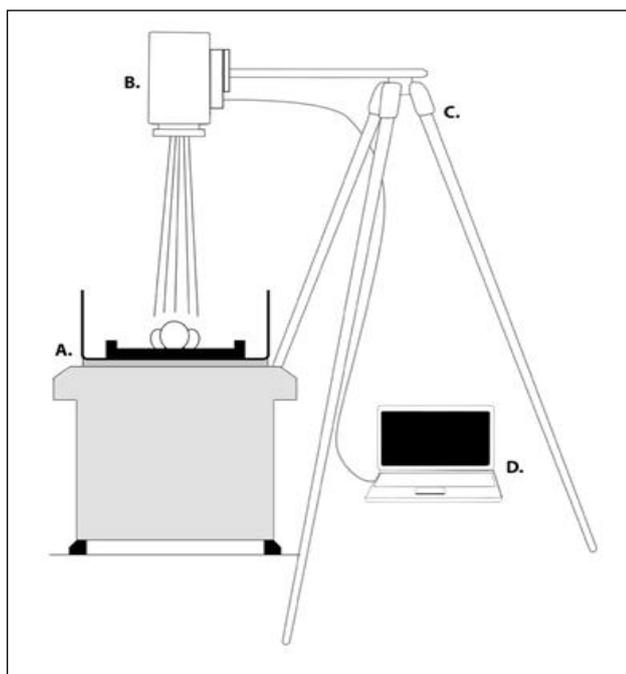
Fig. 1 The schematic set up of the IRT imaging. A: The open radiant warmer. B: The thermal imaging camera. C: The tripod on which the thermal camera was mounted. D: The laptop computer which was connected to the thermal camera during image acquisition for the storage of the thermal images.

The tripod was fully extended in order to allow the camera to be positioned centrally over the incubator at an angle of 90°. For the abdomen to be fully visualized, the infant had to be undressed and the diaper removed. In order to cause as little stress and disturbance to the infant as possible, the thermal imaging scan was timed to coincide with a period where the infant would already be exposed, such as during their daily cares. It was deemed acceptable to allow the infant to acclimatize for two minutes following exposure. A longer period of leaving the infant not covered was considered inappropriate as to the possible stress that may have caused to the infant.

The following patient information was recorded: (i) date of birth, (ii) sex, (iii) birthweight, (iv) gestational age at birth and (v) relevant past medical history. Also at the time of the imaging, (i) current weight, (ii) age, (iii) time of the last oral intake and (iv) time of the last stool passed were noted.

Infants were imaged on the Neonatal Surgical Unit at Sheffield Children's Hospital, Sheffield, United Kingdom. They were housed in Drager Babytherm© open radiant warmers. For the infants included in the study, the heaters of the open radiant warmers were not needed to be operational prior to, and during the thermal scans. The incubator was lowered to the minimum possible height and the tripod and head holding the thermal camera was extended to the maximum possible height of 140 cm. The thermal camera's settings were adjusted to an emissivity of 0.98 (this best matches thermal imaging from the human body) and distance less a one meter (the camera did not have a sub- meter distance setting). It was then positioned centrally over the incubator at an angle of 90°. The distance from the camera lens to the bed was 47 cm and the field of view of the camera was approximately 21 cm. The camera was positioned accordingly to ensure that both abdomen and thorax would be within the field of view. The ambient room temperature and humidity was recorded (mean room temperature: 24.3 °C (standard deviation 1.05 °C), mean room humidity was 34.4 (standard deviation 1.8 across all recordings).

A sequence of thermal images was obtained at a rate of 30 frames per second for 20 seconds. The infant was imaged in the supine position, with all clothing and the diaper removed. If an infant had abdominal dressings or a stoma bag, the thermal scan was scheduled to coincide with the changing or removal of the abdominal coverings.

2.5 Image Processing

As the infants were recorded while awake, there were body movements in some parts the recorded videos. Although a tracking algorithm was incorporated in processing the images (details provided in the following on section), excessive body movement impaired the ability of the tracking algorithm to properly track the region of interest (ROI) across all of the frames. Therefore the videos were viewed and a section that was least affected by body movement or occlusion were selected. This resulted in 7 infants to be represented by 200 images, one infant by 40 images and one infant by 10 images. The stages in processing the images are shown in Figure 2.

Constance Barson,
Reza Saatchi, Prasad Godbole,
Shammi Ramlakhan

Manually select the surgical abdominal region from the first image

Manually select the reference region for the surgical region from the first image

Track the selected region in the each image

Automatically crop the identified region from the images

Represent the region by its mean pixel value

Carry out statistical analysis comparing the surgical region with its reference region

Fig. 2 The stages followed to process the images.

Figure 3 shows the stages of image processing for infant 1. These represent the first image of the recorded video but the remaining images were processed in a similar manner. Figure 3(a) shows the first image for the infant. The image is in RGB colours for visualization, however the images were processed as actual temperature values (indicated by the grey image in Figure 3(b)). Once the first frame was loaded, the irrelevant section of the image, or the background, was excluded to select the relevant area for further processing (Figure 3(c)). This cropping operation was performed by selecting the diagonal co-ordinates of the region of interest and entering them into the Matlab© code. This region was then clustered into five groups using a k-mean clustering algorithm [29] of similar pixel values (Figure 3(d)).
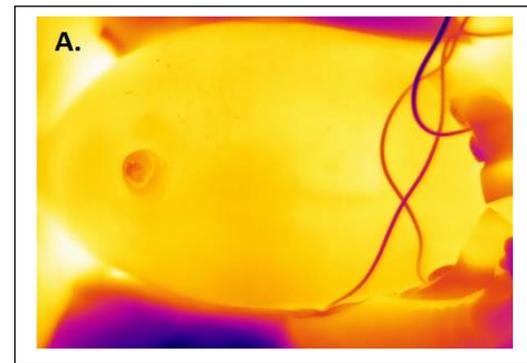
(a)
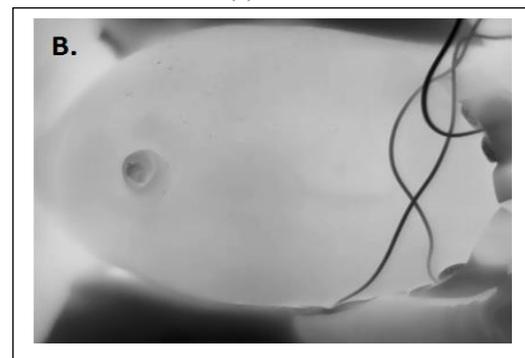
(b)
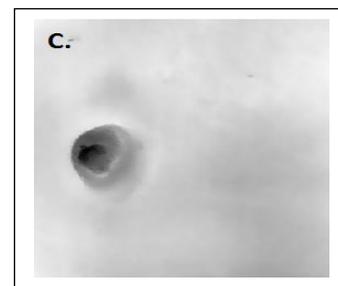
(c)
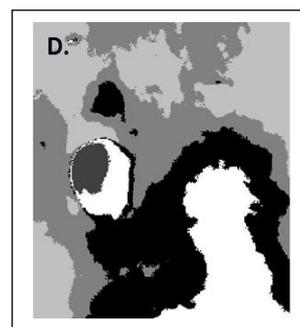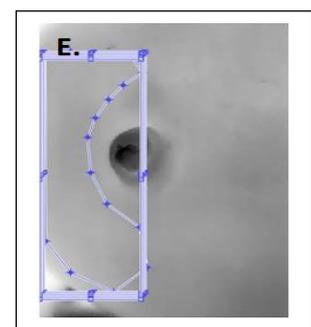
(d)   (e)

Fig. 3 The stages of imaging processing in infant 1. (a) The thermal image in RGB colours. (b). Temperature map. (c) The background is removed, (d) The image is segmented into 5 groups of similar temperature pixels (white represents lowest temperatures) (e) The rectangular region used for tracking and the polygon-selected region of interest (within the rectangular region) used for temperature analysis.

Constance Barson,
Reza Saatchi, Prasad Godbole,
Shammi Ramlakhan

Selection of five groups for clustering was a compromise between specific representation of all distinct regions (requiring larger groupings) and ease of interpretation of the identified regions. The groupings generated a high contrast 'map' that allowed for a better visualization of the temperature distribution in the ROI. Qualitative selection of the inflammatory region and reference region was primarily guided by the knowledge (from the infant's surgical notes) of the where the area of inflammation resulting from surgical intervention was most likely to be. Visual observation of the temperature groupings through clustering in the related region allowed further refinement of the region. To select the ROI, the first image of the recorded video was displayed on the computer screen together with the clustering image. The polygon function in Matlab© was evoked in the Matlab© code. This changed the cursor pointer to a cross pointer. Using the mouse, the cross pointer was positioned on the starting point in the region to be selected and clicked. The selected position appeared as a dot on the image. By moving the cursor gradually forward in the path of the desired region (each time observing the clustering image) and clicking the mouse at each move, a complete closed polygon was specified. The places where the mouse was clicked are shown as jointed dots as shown in Figure 3(e).

When selecting the inflammatory and reference regions, care was taken to ensure that they did not include the navel region or surgical scar as their tissue could have a different temperature profile to the healthy tissue.

Once the ROI from the first image was selected, the region was tracked using the template matching tracking algorithm [30] in the following images. The tracking algorithm used normalized cross correlation (NCC) to find a region in each following image that was most similar to the selected ROI from the first image. NCC has been found to be more accurate than other similarity measurement methods such as the sum of absolute difference and sum of squared difference [31]. For the purpose of tracing, the rectangular region shown in Figure 3(e) was used. This enclosed the ROI and its coordinates were automatically determined from the coordinates of the polygon through the Matlab© code.

Tracking was included to deal with body movements and to align the selected ROI across all images within the video. Once the selected ROI was identified in each image of the recorded video for the infant, the pixel values (representing temperature) contained within the polygon were averaged. These were further averaged across all images, representing the video, to produce a single mean temperature for the region.

The reference ROI was cropped in the similar manner to the surgical region except its abdominal location was carefully selected to be as further away as possible from the selected surgical region, to be approximately the same size and avoided areas that could have unduly affected its temperature.

2.6   Data Analysis

In order to compare the temperature of the surgical ROI and its reference section across the 9 infants included in the analysis, the statistics of the temperatures were obtained. The statistics were expressed in terms of mean, mode, median, standard deviation, minimum, maximum, skewness and kurtosis variables. These 8 variables were analyzed using IBM SPSS© Statistics 25 for Mac© [32] to establish whether they would allow differentiation between inflammatory and their associated reference regions. Tests of normality were performed on the variables and their distributions were visualized with quantile-quantile (Q-Q) plots. A Q-Q plot is a scatter diagram of two sets of quantiles against each other. If the quantiles are from the same distribution type, the points form a roughly diagonal line. The paired samples t-test (a parametric test) was performed on the normally distributed statistical variables to determine whether the difference in the means between the inflammatory and reference regions was significantly different from zero. For statistical parameters which were not normally distributed, the equivalent non-parametric test, the Wilcoxon-Signed Ranks test was applied.

In order to investigate the possible influence that the infants' demographic information may have on skewness, multiple linear regression (MLR) was conducted. MLR assumes a linear relationship between the dependent and independent variables, multivariate normality, homoscedasticity (this assumes the variance around the regression line is the same for all predictor variable values) and no multicollinearity (occurrence of high inter-correlations among independent variables). As the sample size was small, 3 separate analyses were conducted referred to as models 1 to 3, with the dependent variable as skewness (this feature was most discriminatory as explained in a later section). The models were:
- Model 1: used the current age and gestational age at birth as independent variables.
- Model 2: used the current weight and birthweight as independent variables.

- Model 3: used the last intake and last stool prior to imaging as independent variables.

The MLR's assumption of a linear relationship between the dependent and independent variables were not met for the models. Therefore, the data were transformed, using a non-linear transformation, by taking the square root of each observation in the independent variables age at the time of imaging, gestational age at birth, weight at the time of imaging, birthweight, last stool prior to imaging and last oral intake prior to imaging. The MLR conformance was checked again by considering the following plots.

- Histogram of the standardized residuals for the difference in skewness, with the square root of age at the time of imaging and age at birth as covariates. Residual is the difference between the data response and fitted response and standardized residual is the defined as the residual divided by its standard deviation.
- Probability-probability (P-P) plot of the standardized residuals for the difference in skewness in MLR, with the square root of age at the time of imaging and age at birth as covariates. P-P plot can indicate the extent two data sets agree. It is a plot of two cumulative density functions against each other. The points in the plot should approximately fall onto a diagonal line, if the specified distribution is correct.
- Scatterplot of the standardized residuals for the difference in skewness in MLR, with the square root of age at the time of imaging and age at birth as covariates.

3 Results

3.1 Summary Statistics of the Study Population

The overall summary statistics of the demographics for the 9 infants included in the analysis are provided in Table 1.

As the sample size was small, the demographic information for the infants individually and their surgical conditions are included Appendix A.

Table 1 Summary statistics of the demographics for the infants. Gestational age, weight at the time of imaging, last stool before imaging and room temperature are given as median (interquartile range, IQR), sex (expressed numerically) and the rest are given as mean ($\pm$ SD).

Variables		Surgical group (n=9)
Age at time of imaging (days)		58 (21 − 83)
Gestational age (days)		227 (213.5 − 259)
Weight at time of imaging (kg)		2.65 (2.45 − 3.15)
Birthweight (kg)		1.89 ($\pm$ 0.636)
Sex	Male	3
	Female	6
Last oral intake before imaging (minutes)		106 ($\pm$ 90.4)
Last stool before imaging (minutes)		10 (10 − 25)
Room temperature (ºC)		24.9 (23 − 25.2)
Room humidity		24.3 ($\pm$1.06)

3.2 Data Visualization

The boxplots of the differences between the inflammatory region and its reference in each variable are provided in Figure 4.
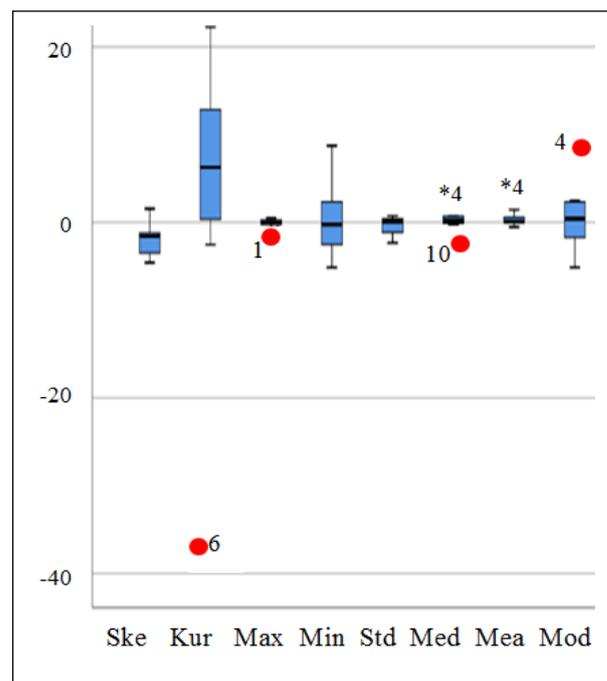
Fig. 4 Boxplots of the difference in each variable. Outlier values are represented by stars and extreme outlier values are represented by red dots. The infant ID is indicated next to each outlier (key:- Ske: Skewness, Kur: Kurtosis, Max: Maximum, Min: Minimum Std: Standard Deviation, Med: Median, Mea: Mean, Mod: Mode).

The largest spread for the difference was associated with kurtosis while maximum, median and mean showed least difference.

The boxplots illustrated in Figures 5 - 12 compare the distributions of the 8 variables for the two different regions. The medians of all variables, apart from skewness and kurtosis, are very similar. Outlier values are represented by black stars and extreme outlier values are represented by red dots. The infant ID is indicated next to each outlier.
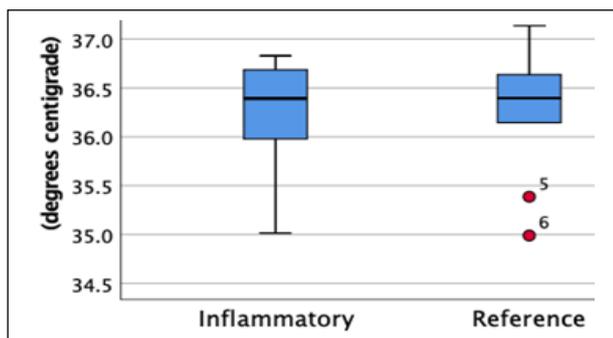
Fig. 5 Boxplots of the inflammatory and reference regions for the maximum.
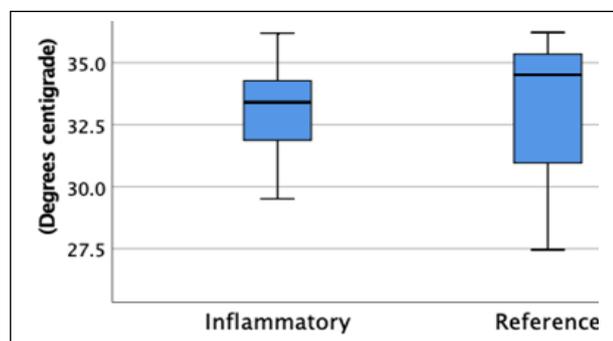
Fig. 6 Boxplots of the inflammatory and reference regions for the minimum.
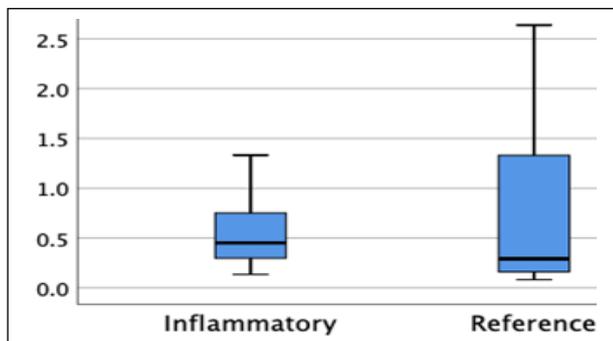
Fig. 7 Boxplots of the inflammatory and reference regions for the standard deviation.
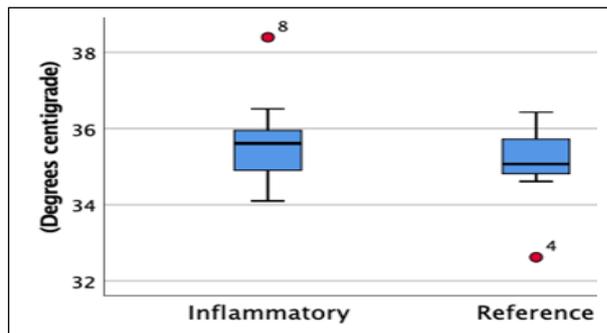
Fig. 8 Boxplots of the inflammatory and reference regions for the median.
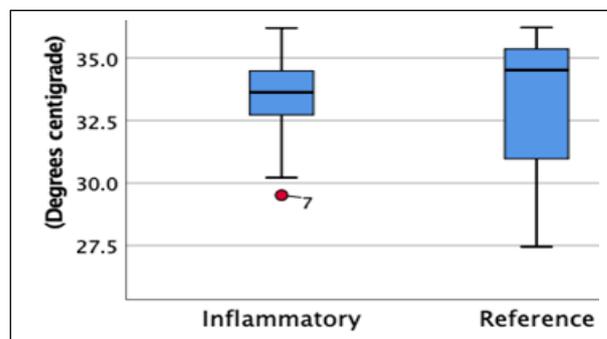
Fig. 9 Boxplots of the inflammatory and reference regions for the mode.
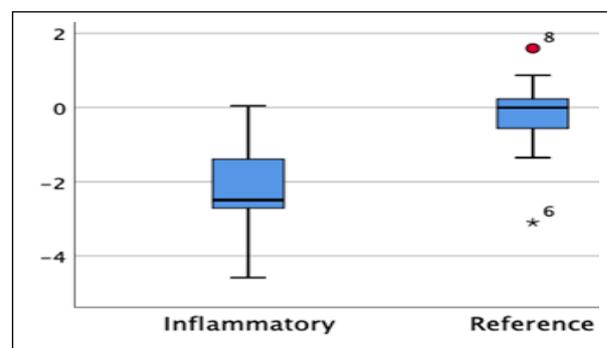
Fig. 10 Boxplots of the inflammatory and reference regions for the mode skewness.
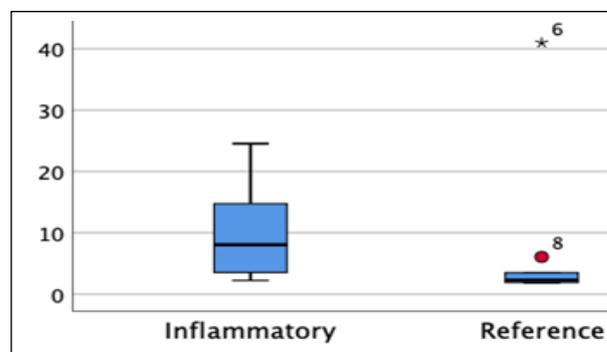
Fig. 11 Boxplots of the inflammatory and reference regions for the mode kurtosis.

Constance Barson,
Reza Saatchi, Prasad Godbole,
Shammi Ramlakhan
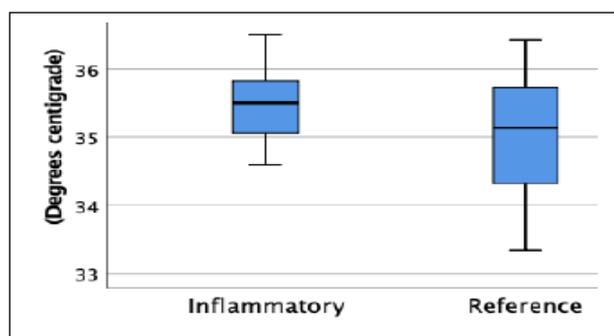
Fig. 12 Boxplots of the inflammatory and reference regions for the mean.

The spread of the data varied for both regions of the 8 variables. The reference region had a higher interquartile range for the minimum, standard deviation, mean and mode whereas the spread of data was larger in the inflammatory region for the maximum, skewness, kurtosis and median. The only variables which did not exhibit significant skew were the maximum and mean. A number of outliers were visualized, including the extreme outlier in the kurtosis of the reference region. Owing to the small sample size and variability in surgical condition, the outliers were not removed from the data set as the outlier value may be representative of a particular surgical condition.

3.3   Statistical Tests

In order to quantify the relationship between the inflammatory and reference regions, the Spearman's rank-order correlation test was performed as it is less sensitive to outliers than Pearson's correlation. The results of the Spearman's rank-order correlation test are shown in Table 2.

Table 2 The Spearman's rank-order correlation test for examining the relationship between the inflammatory and reference regions for all variables.

Variable	Correlation coefficient	Significance (α=0.05)
Maximum	0.367	0.332
Minimum	-0.117	0.765
Mean	0.283	0.460
Standard deviation	-0.217	0.576
Median	-0.167	0.668
Mode	0.117	0.765
Skewness	0.117	0.765
Kurtosis	0.183	0.637

The null hypothesis for the Spearman correlation is that there is no monotonic relationship between the inflammatory region and reference region of each variable. All variables showed a weak

correlation and none showed the correlation to be statistically significant.

Prior to determining which variable provided statistically significant measure for distinguishing between the inflammatory and reference regions, the Shapiro-Wilk test was applied to determine whether the data were from a normal distribution. The Shapiro-Wilk test was chosen over the Kolmogorov-Smirnov test as it exhibits greater sensitivity for small sample sizes (<50 samples). In the case of a normal distribution the Shapiro-Wilk statistics statistic (W) is close to 1. If its p-value is less than the chosen alpha level (the value for this level was 0.05), the null hypothesis that the data are from a normal distribution is rejected. The summary statistics of the inflammatory and reference regions are presented in Table 3. Any variable which was found not to be normally distributed was expressed as the median (accompanied with interquartile range, IQR) and any variable found to be normally distributed was expressed as the mean (accompanied with standard deviation, std).

Table 3 Summary statistics for the inflammatory and reference regions of all statistical parameters for the 9 infants. Parentheses either indicate the IQR range or standard deviation.

Variable	Region	Variable value
Maximum	Inflammatory	36.4 (35.6 – 36.7)
	Reference	36.4 (35.7 – 36.8)
Minimum	Inflammatory	32.9 (2.12)
	Reference	32.8 (3.28)
Mean	Inflammatory	35.5 (35.0 – 35.9)
	Reference	35.1 (34.0 – 35.9)
Standard deviation	Inflammatory	0.45 (0.24 – 0.81)
	Reference	0.29 (0.14 – 1.9)
Median	Inflammatory	35.6 (34.8 – 36.2)
	Reference	35.1 (34.7 – 35.8)
Mode	Inflammatory	33.6 (31.5 – 35.2)
	Reference	34.5 (29.7 – 35.4)
Skewness	Inflammatory	-2.09 (1.54)
	Reference	-0.272 (1.35)
Kurtosis	Inflammatory	8.06 (3.48 – 16.7)
	Reference	2.30 (1.90 – 4.79)

Statistical tests of significance were carried out to establish whether the difference in the temperature of the inflammatory and reference regions were significant. For variables found to be from a normal distribution, paired sample t-test was used. This test determined whether the difference in the means between paired observations was significantly different from zero. The null hypothesis was that the

difference in the means between the two regions for the variables was equal to zero. A statistically significant result ($p < 0.05$) implies that a particular variable can be used to distinguish between the two regions. The assumptions for a paired samples t-test are that the dependent variable is continuous, the samples are related, the difference between the paired samples of data are approximately normally distributed and that there are no outliers in the difference between the paired samples. A detailed discussion of issues related to t-test is provided in reference [33]. The assumptions were met for the difference in skewness and difference in minimum for the inflammatory and reference regions. The results for the associated paired samples t-test are shown in Table 4.

Table 4 Paired samples t-tests for the temperature difference of inflammatory and reference regions (significance level = 0.05).

Variable	Mean difference (95% CI)	p- values
Skewness	-1.816 (-3.293, -0.339)	0.022
Minimum	0.157 (-3.018, 3.331)	0.912

There was a significant mean difference between the skewness of the inflammatory region and skewness of the reference region (p=0.022). On average, the inflammatory regions were more negatively skewed by 1.816 as compared to the reference regions. The difference in mean between the minimum (temperature) of the inflammatory region and minimum of the reference region was found not to be statistically significant (p = 0.912).

For the remaining variables that did not conform to a normal distribution, the Wilcoxon signed-rank test, a non-parametric test, was conducted. The Wilcoxon signed-rank test computed the differences between pairs of samples and compared the sample median against a hypothetical median. The null hypothesis was that the median difference between pairs of samples was zero. None of the remaining variables showed significant difference. Thus the only significant predictor that distinguished between the inflammatory region and the reference region was skewness. For this variable, further details of its characteristics are provided next. Figure 13 shows the Q-Q plot and Figure 14 shows the histogram for the difference of the inflammatory and reference regions for the skewness. The Q-Q plot compared the quantiles of the difference in the skewness of the inflammatory and reference regions versus

theoretical quantiles from a normal distribution. The quantile-quantile (Q-Q) plot appeared to mostly fall on, or very close to the reference (linear) line, indicating a close to normal distribution. This is further evidenced by the histogram plot of the difference that indicated the distribution was approximately normally distributed.
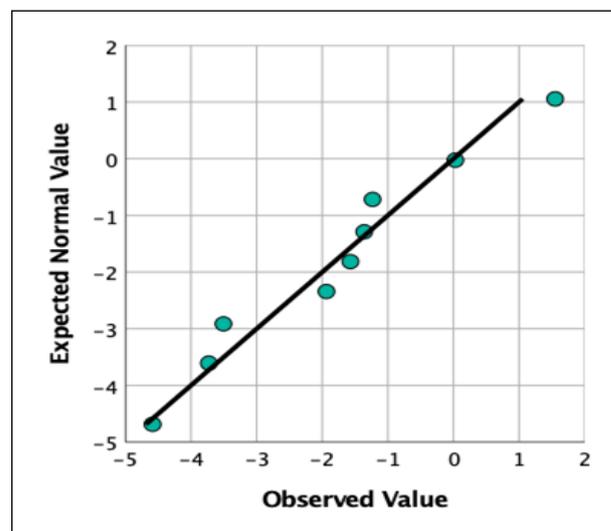
Fig. 13 The Q-Q plot for the difference of the inflammatory and reference regions for the skewness.
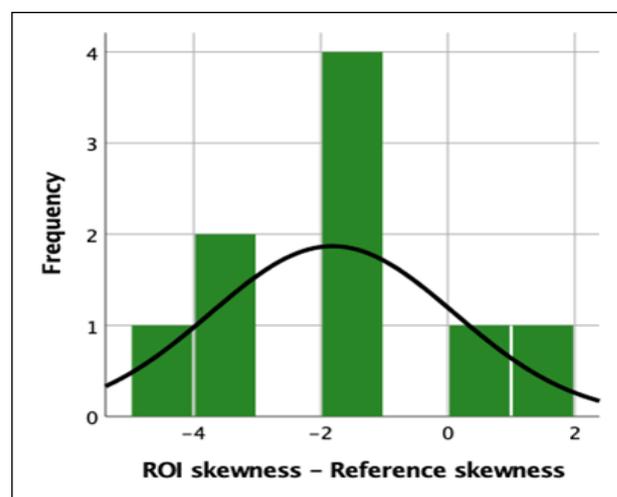
Fig. 14 The histogram for the difference of the inflammatory and reference regions for the skewness.

As indicated in the methodology, in order to investigate the possible influence that the infants' demographic information may have on skewness, multiple linear regression (MLR) was conducted based on three models referred to earlier. As an illustration, the conformity tests for model 1 are provided (the same tests were performed for models

2 and 3). Figure 15 shows the histogram of the standardized residuals for the difference in skewness, with the square root of age at the time of imaging and age at birth as covariates.
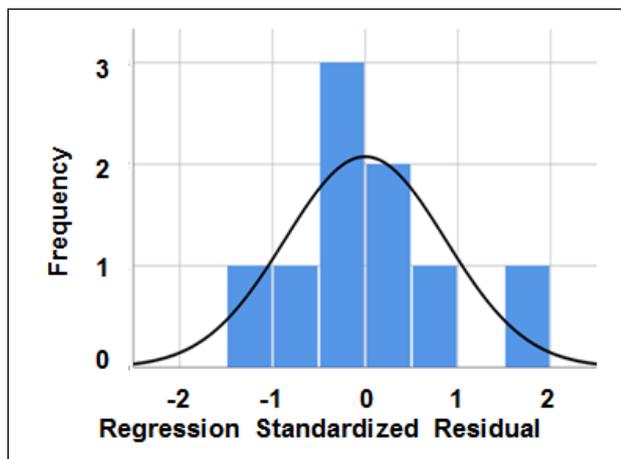
Fig. 15 Test associated with MLR for model 1. (a) Histogram of the standardized residuals for the difference in skewness, with the square root of age at the time of imaging and age at birth as covariates.

This histogram was close to normal distribution. The assumption of normally distributed residuals was therefore met. Figure 16 shows the probability-probability (P-P) plot of the standardized residuals for the difference in skewness in MLR, with the square root of age at the time of imaging and age at birth as covariates.
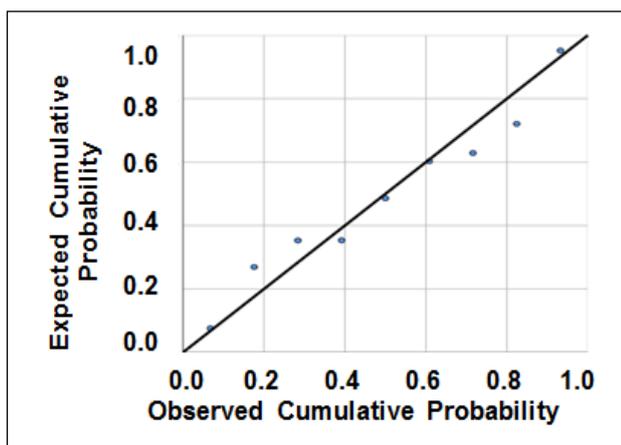
Fig. 16 Test associated with MLR for model 1 P-P plot of the standardized residuals for the difference in skewness in MLR, with the square root of age at the time of imaging and age at birth as covariates.

A symmetrical s-shaped P-P plot was observed, indicating that the model overestimates the values of skewness as much as it underestimates them. As the s-shape is symmetrical, the residuals were approximately symmetrically distributed.

Figure 17 shows the scatterplot of the standardized residuals for the difference in skewness in MLR, with the square root of age at the time of imaging and age at birth as covariates. There was no evidence of heteroscedasticity (i.e. unequal variance of the data along the regression line).
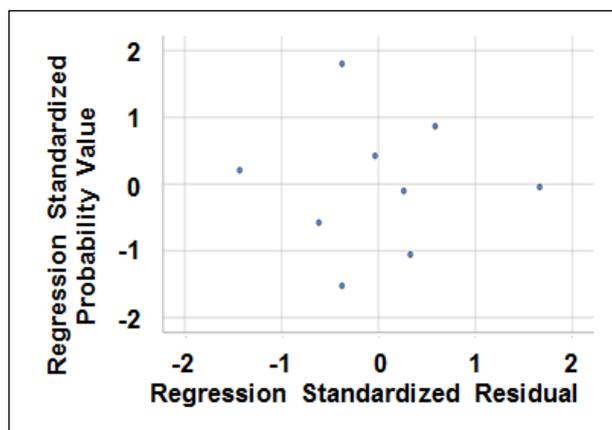
Fig. 17 Test associated with MLR for model 1 Scatterplot of the standardized residuals for the difference in skewness in MLR, with the square root of age at the time of imaging and age at birth as covariates.

Tables 5 and 6 show the results for multiple linear regression conducted to determine the strength of the effect of patient demographic information on the difference in skewness.

Tables 5 The R^2 and adjusted R^2 values for the MLR models.

Models	Independent variable	R^2	Adjusted R^2
1	Age at imaging	0.152	-0.130
	Gestational age at birth		
2	Weight at imaging	0.187	-0.084
	Birthweight		
3	Last stool prior to imaging	0.224	-0.034

Tables 6 The probability (p) values for the MLR models.

Models	Unstandardized coefficient (CI)	p-value
1	-0.262 (-1.64, 1.12)	0.659
	-1.32 (-5.11, 2.48)	0.428
2	-3.15 (-12.7, 6.37)	0.449
	-2.00 (-8.98, 4.97)	0.509
3	-0.235 (-0.734, 0.265)	0.295
	-0.058 (-0.389, 0.272)	0.681

The dependent variable for MLR (Tables 5 and 6) was for the difference in skewness of the inflammatory and reference regions. A significance level of 0.05 was used for this test. The independent variables in all models were transformed by taking the square root of each observation in order to satisfy the assumptions for multiple linear regression. The adjusted R^2 (a statistical measure indicating the extent of closeness of the data to the fitted regression line) values were close to zero for all models, indicating that the independent variables were poor predictors for the difference in skewness.

The R^2 value in model 1 indicated that 15.2% of the variance in the difference in skewness may be attributed to age at the time of imaging and gestational age. For model 2, 18.7% of the variance in difference in skewness may be explained by the weight at imaging and the birthweight. For model 3, 22.4% of the variance in difference in skewness may be attributed to last stool prior to imaging and last oral intake prior to imaging. The overall regression model for models 1-3 was not a good fit for the data as signified by the F-test results: Model 1: F = 0.539, p = 0. 609; Model 2: F = 0.690, p = 0.538; Model 3: F = 0.868, p = 0.467. The F-test for linear regression indicates whether any of the independent variables in MLR models are significant (unlike the t-test that considers coefficients of individual variables on their own for significance, F-test considers coefficients of multiple variables simultaneously). As the significance value for each independent variable in all models was greater than 0.05, the null hypothesis that the coefficients were statistically significantly different from zero could not be rejected. Thus, the predictors: age at imaging, gestational age at birth, weight at imaging, birthweight, last stool prior to imaging and last oral intake prior to imaging did not have a significant effect on the difference in skewness.

4   Discussion

In our study, it would not be feasible to select a standardized ROI based on anatomical landmarks for each infant as the surgical conditions differed considerably. ROI selection was therefore based on primarily on the knowledge of surgery location and the appearance of the segmented image, which clustered groups of pixels of similar temperature together, and thus indicating the presence of features. Although this technique received guidance by the segmented image, it is likely to be influenced by interpreter bias. In further studies, if this approach was used, it may be preferable to use several trained thermographers to identify the inflammatory and reference regions.

The tests of statistical significance determined that of the 8 statistical variables extracted from the thermal images, skewness was the only statistically significant variable (p = 0.022). It was found that on average, the skewness value of the inflammatory regions was less than the reference regions (mean difference = -1.816 (95% CI: -3.293, -0.339)). This indicated that for inflammatory regions, more pixel temperature values fell towards the higher end of the scale and less pixel temperature values fell towards the lower end of the scale.

In order to determine whether the infants' demographic information had any effect on the difference in skewness, MLR was conducted. Three separate analyses were conducted due to the small sample size and for each of the three models, the adjusted R^2 value was found to be close to zero, indicating that overall regression models were not a good fit for the data and that the predictors used were poor predictors for the difference in skewness. This was also confirmed by the lack of statistical significance of the unstandardized coefficients for each independent variable. The independent variables used in the analysis include: age at imaging, gestational age at birth, weight at imaging, birthweight, last stool prior to imaging and last oral intake prior to imaging. The results for the MLR analysis need to be put in the context of small sample size. Of the three studies referred to in this paper, which have explored the use of thermography on infants with NEC, i.e. [11]. [15], [34], only the study reported in [15] had incorporated the patient demographic information into their analysis. They chose to include clinical features such as the sex, weight at birth, the current weight and corrected age into an algorithm, which was trained to identify features on the thermograms from the NEC cohort.

This study, by identifying a significant discriminant, i.e. skewness, has shown that the

technique could be a feasible method for distinguishing between healthy and inflamed tissue. Skewness represents asymmetry in relation to the symmetrical normal distribution. [35]. It therefore characterizes the manner temperatures from the ROI relate to each other. The finding of this study indicates that, this temperature relationship is more important in discriminating between inflammatory region and its reference than the magnitude of the temperatures that can be quantified by features such as minimum, maximum, median, mode and mean.

This study was a feasibility study; future research would aim to explore the issues in the next section.

5   Direction of Future Research

A number of improvements could be made following identification of the limitations of this study. A larger number of infants, specifically premature babies with NEC, premature age-matched controls without abdominal pathology and healthy term babies would be valuable to allow for a more thorough statistical evaluation. The inclusion of the premature aged-matched infants and healthy term infants would allow for the comparison of thermal patterns with those obtained from NEC babies with a view to identify significant differences. Increasing the cohort size would allow for greater accuracy of results as it would be possible to identify the individual variation in skin temperature across a range of infants of different ages, races and genders and how this influences the thermal pattern.

Automating the selection of ROI would reduce possible bias and helps with consistency of segmentation operation.

Artificial intelligence (AI) methods such as neural networks and fuzzy logic could be trained to assist with differentiation of thermal patterns thus reducing the expertise needed in their interpretation. AI techniques have proved effective in medical diagnosis and interpretation of medical data [36], [37]. The advances in the AI method of deep learning [39] will open new analysis opportunities in in interpreting the thermal images.

A further area of future exploration could be the inclusion of active dynamic thermal imaging. Active dynamic thermal imaging involves the application of a hot or cold external thermal stimulus, known as a thermal stress test, to a subject for a defined period of time [39]. The use of a thermal stress prior to the acquisition of thermal images allows for the identification of areas of cooling, warming or

thermal oscillation and represents an alternate method for qualitatively selecting a region of interest without the use of clustering. Following the application of a hot or cold stimulus to the abdomen, the time taken for thermal recovery can be observed and recorded. A possible advantage of active dynamic thermal imaging is that the use of a thermal stress test may enhance the thermal contrast of the region of interest and allow for the observation of the temporal evolution of thermal patterns [39]. As NEC is thought to have an ischaemic component in its pathophysiology, it may be feasible that the inclusion of a thermal stress test in the thermal imaging procedure for babies with NEC might enhance the area of pathology and the time taken for thermal recovery may potentially indicate the severity of disease [40], [41].

6   Conclusion

Thermal imaging techniques were developed to detect the temperature profile of the abdomen in surgical infants for distinguishing between healthy tissue and inflammatory tissue. The study was conducted as a proof of concept study to determine if dynamic passive thermal imaging would provide significant discrimination.

The mean difference between the skewness of the inflammatory region and skewness of the reference region were significantly different (p=0.022). On average, the inflammatory regions were more negatively skewed by 1.816 as compared to its abdominal reference regions.

The results of this study indicated that inflammatory regions, such as those found in infants following surgery, would have a higher temperature distribution profile than the surrounding skin. The early findings of this feasibility study need to be confirmed using a larger population of infants.

The strengths of this work included: (i) using passive dynamic infrared thermal imaging, as compared with single imaging, that further assisted with noise reduction through temporal averaging, (ii) use of clustering in identification of the region of interest, (ii) inclusion of tracking algorithm to deal with infant's body movements, (iii) each infant acting as their own control and (v) detailed statistical analysis to identify the most suitable discriminatory variable.

References

1. Owen, R. and Ramlakhan, S., Infrared thermography in paediatrics: a narrative review of clinical use, *BMJ Paediatrics Open*, 2017, doi:10.1136/bmjpo-2017-000080, https://bmjpaedsopen.bmj.com/content/bmjpo/1/1/e000080.full.pdf

2. Knobel, R.B., Guenther, B.D. and Rice, H.E., Thermoregulation and thermography in neonatal physiology and disease, *Biological Research For Nursing*, Vol. 13, No.3, 2011, pp. 274–282.

3. Anbar, M., Objective assessment of clinical computerized thermal images, In: Loew, M.H. Editor, *International Society for Optics and Photonics*, 1991, pp. 479–484.

4. Elphick, H.E., Alkali, A.H., Kingshott, R.K., Burke, D. and Saatchi R., Exploratory study to evaluate respiratory rate using a thermal imaging camera, *Respiration*, Vol. 97, No. 3, 2019, pp. 205-212.

5. Alkali, A.H., Saatchi, R., Elphick, H. and Burke, D., Thermal image processing for real-time non-contact respiration rate monitoring, *IET Circuits, Devices and Systems*, Vol 11, No. 2, 2017, pp. 142-148.

6. Priego Quesada, J.I., Sanchis-Sánchez, E., Salvador, R., Codoñer-Franch, P., Martín-Guerrero, J., Vergara-Hernández, C., Blasco, J., Ballester, E., Sanchis, E., González-Peña, R. and Cibrian, R., Clinical applications: infrared thermal diagnosis of orthopaedic injuries in childhood, In Innovative Research in Thermal Imaging for Biology and Medicine, Editors Vardasca, R. and Mendes, J.G, *IGI Global*, Chapter 3, 2017, pp. 55–78, ISSN: 2327-9354; eISSN: 2327-9370.

7. De Salis, A.F., Saatchi, R. and Dimitri, P., Evaluation of high resolution thermal imaging to determine the effect of vertebral fractures on associated skin surface temperature in children with osteogenesis imperfecta, *Medical & Biological Engineering & Computing,* Vol.56, No.9, 2018, pp. 1633-1643.

http://doi.org/10.1007/s11517-018-1806-3.

8. Owen, R., Ramlakhan, S., Saatchi, R. and Burke, D., Development of a high-resolution infrared thermographic imaging method as a diagnostic tool for acute undifferentiated limp in young children, *Medical & Biological Engineering & Computing*, Vol.56, No.6, 2018, pp. 1115-1125. http://doi.org/10.1007/s11517-017-1749-0.

9. Reed, C., Saatchi, R., Burke, D. and Ramlakhan, S., Infrared thermal imaging as a screening tool for paediatric wrist fractures, *Medical and Biological Engineering and Computing*, 2020, http://doi.org/10.1007/s11517-020-02167-z

10. Nwaizu, H., Saatchi, R., Hawley, D.P. and Ward, O., Thermal and visual imaging to assist with juvenile idiopathic arthritis examination of the knees, *Technologies*, Vol.8 (30), 2020, pp. 1-19.

11. Clark, R.P. and Stotherst, J.K., Neonatal skin temperature distribution using infra-red colour thermography, J. Physiol., Vol. 302, 1980, pp. 323–333.

12. Christidis, I., Zotter, H., Rosegger, H., Engele, H., Kurz, R. and Kerbl, R., "Infrared thermography in newborns: the first hour after birth," Gynäkologisch-geburtshilfliche Rundschau, vol. 43, no. 1, pp.:31–35, 2003.

13. Saxena, A.K., Willital, G.H., Infrared thermography: Experience from a decade of pediatric imaging, *European Journal of Pediatrics*, Vol. 167, No. 7, 2008, pp. 757–764.

14. Rice, H.E., Hollingsworth, C.L., Bradsher, E., Danko, M.E., Crosby, S.M., Goldberg, R.N., Tanaka, D.T., Knobel, R.B. and Carolina, N., Infrared thermal imaging (thermography) of the abdomen in extremely low birthweight infants, *J Surg Rad*., Vol 1, No. 2, 2010, pp. 82–89.

15. Herry, C.L., Frize, M. and Bariciak, E., Assessment of abdominal skin temperature change in premature newborns with NEC

compared to healthy controls, *In 5th European Conference of the International Federation for Medical and Biological Engineering*, 2011, pp. 191–194. http://link.springer.com/10.1007/978-3-642-23508-5_51.

16. Knobel-Dail, R.B., Holditch-Davis, D., Sloane, R., Guenther, B.D. and Katz, L.M., Body temperature in premature infants during the first week of life: Exploration using infrared thermal imaging, *Journal of Thermal Biology*, Vol. 69, 2017, pp. 118–123.

17. Gregory, K.E., Deforge, C.E., Natale, K.M., Phillips, M. and Van Marter, L.J., Necrotizing enterocolitis in the premature infant: neonatal nursing assessment, disease pathogenesis, and clinical presentation, *Advances in Neonatal Care: Official Journal of the National Association of Neonatal Nurses*, Vol. 11, No. 3, 2011, pp. 155-164.

18. Jones, B.F., A reappraisal of the use of infrared thermal image analysis in medicine, *IEEE Transactions on Medical Imaging*, Vol. 17, No. 6, 1998, pp.1019–1027.

19. Pomerance, J.J., Lieberman, R.L. and Ukrainski, C.T., Neonatal thermography, *Pediatrics*, Vol. 59, No. 3, 1977, pp. 345–51.

20. Neu, J., Modi, N. and Caplan, M., Necrotizing enterocolitis comes in different forms: historical perspectives and defining the disease, *Seminars in Fetal and Neonatal Medicine*, Vol. 23, No. 6, 2018, pp. 370-373, doi: 10.1016/j.siny.2018.07.004.

21. Sharma, R. and Hudak, M., A clinical perspective of necrotizing enterocolitis: past, present, and future, *Clinics in Perinatology*, Vol. 40, No. 1, 2013, pp. 27–51.

22. Anderson-Berry, A.L., Bellig, L.L. and Ohning, B.L., Neonatal sepsis: background, pathophysiology, etiology, *Neonatal Sepsis*. [Online], 2015. https://emedicine.medscape.com/article/978352-overview.

23. Epelman, M., Daneman, A., Navarro, O.M., Morag, I., Moore, A.M., Kim, J.H., Faingold, R., Taylor, G. and Gerstle, J.T., Necrotizing enterocolitis: review of state-of-the-art imaging findings with pathologic correlation, *RadioGraphics*, Vol. 27, No. 2, 2007, pp. 285–305.

24. Bohnhorst, B., Usefulness of abdominal ultrasound in diagnosing necrotising enterocolitis, *Archives of Disease in Childhood: Fetal and Neonatal Edition*, Vol. 98, No. 5, 2013, pp. 445–450.

25. Good Clinical Practice (GCP), *National Institute for Health Research*, https://www.nihr.ac.uk/health-and-care-professionals/learning-and-support/good-clinical-practice.htm.

26. Güney, L.H., Araz, C., Beyazpınar, D.S., Arda, İ.S., Arslan, E.E. and Hiçsönmez, A., Abdominal problems in children with congenital cardiovascular abnormalities, *Balkan Medical Journal*, Vol. 32, No. 3, 2015, pp. 285–290.

27. Waldron, S. and MacKinnon, R., Neonatal thermoregulation, *Infant*, Vol. 3, No. 3, 2007, pp. 101–104.

28. Matlab©, Version, *Mathworks Inc, Cambridge*, 2017.

29. Shan, P, Image segmentation method based on K-mean algorithm, *Eurasip Journal on Image and Video Processing*, 2018 https://doi.org/10.1186/s13640-018-0322-6.

30. Brunelli, R. Template matching techniques in computer vision: theory and practice, *Wiley.*, 2009, ISBN: 978-0-470-51706-2.

31. Munsayac, F.E.T. Alonzo, L.M.B. Lindo, D.E.G. Baldovino, R.G. Bugtai, N.T., Implementation of a normalized cross-correlation coefficient-based template matching algorithm in number system conversion. *In Proceedings of the IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*, Manila, Philippines, 1–3 December 2017; Corpus

39411265, doi:10.1109/HNICEM.2017.8269520.

32. SPSS© Inc., Chicago, I.L., USA. https://www.ibm.com/analytics/spss-statistics-software.

33. Derrick, B., Russ, B., Toher, D. and White, P., Test Statistics for the Comparison of Means for Two Samples That Include Both Paired and Independent Observations, Journal of Modern Applied Statistical Methods, Vol. 16, No. 1, 2017, pp. 137-157, doi: 10.22237/jmasm/1493597280.

34. Ntonfo, G.M.K., Frize, M. and Bariciak, E., Detection of necrotizing enterocolitis in newborns using abdominal thermal signature analysis, In: *2015 IEEE International Symposium on Medical Measurements and Applications (MeMeA) Proceedings [Online]*. 2015, pp. 36–39. http://ieeexplore.ieee.org/document/7145168/.

35. Doane, D.P. and Seward, L.E., Measuring Skewness: A Forgotten Statistic?, *Journal of Statistics Education*, Vol.19, No. 2, 2011, pp.1-18.

36. Saatchi, R., Oke, S., Allan, E., Jervis, B. and Huddon, N., Signal processing of the contingent negative variation in schizophrenia using multilayer perceptrons and predictive statistical diagnosis, *IEE Proceedings. Science Measurement and Technology*, Vol. 142, No. 4, 1995, pp. 296-277. http://doi.org/10.1049/ip-smt:19951838

37. Saatchi, R., Single-trial lambda wave identification using a fuzzy inference system and predictive statistical diagnosis. *J Neural Eng*, Vol.1, No.1, 2004, pp. 21-31.

38. Wang, F., Casalino, L.P. and Khullar, D., Deep learning in medicine—promise, progress, and challenges. *JAMA Internal Medicine*. Vol.179, No. 3, 2019, pp.293-294, doi:10.1001/jamainternmed.2018.7117.

39. Baek, Y.S., Kim, J., Han, G. and Oh, C.H., 2018. Application of dynamic thermal imaging in a photocarcinogenesis mouse model. *International Journal of Hyperthermia*. Vol.34, No.7, 2018, pp.961–968.

40. Choi, Y.Y., Necrotizing enterocolitis in newborns: Update in pathophysiology and newly emerging therapeutic strategies. *Korean Journal of Pediatrics*. Vol. 57, No. 12, 2014, pp.505–513.

41. Dollberg, S., Lusky, A. and Reichman, B., Patent ductus arteriosus, indomethacin and necrotizing enterocolitis in very low birth weight infants: a population-based dtudy, *Journal of Pediatric Gastroenterology and Nutrition*. Vol.40, No.2, 2005, pp.184–188.

Contribution of Individual Authors

All authors have contributed to the work and the write up of the article.

Constance Barson,
Reza Saatchi, Prasad Godbole,
Shammi Ramlakhan

Appendix A   Infants' demographic information.

ID	Sex	Gestational age at birth (weeks)	Corrected gestational age at time of imaging (weeks)	Weight at time of imaging (kg)	Birth-weight (kg)	Past medical history	Last oral intake prior to imaging (minutes)	Last stool prior to imaging (minutes)	Room temperature (ºC)	Room humidity
1	F	37+0	39+5	2.40	2.24	Correction of jejunal atresia 18 days prior to imaging	5	30	25.0	33
2	F	35+5	46	4.45	2.40	Correction of mid-jejunal atresia and dilation and mid-colonic stenosis by jejunostomy, colectomy and end to end anastomosis of colon to colon	180	180	23.1	33
3	M	32+1	40+3	2.97	1.78	NEC with perforation Laparotomy and ileostomy formation	60	10	25.2	32
4	F	28+6	42+2	3.25	1.18	NEC confirmed via laparotomy Stoma formed 35cm from duodenojejunal flexure	10	10	25.1	31
5	F	37+0	38+2	2.28	2.47	Anorectal malformation, imperforate anus and recto-vestibular malformation Colostomy involving sigmoid to descending colon 8 days prior to imaging	180	5	22.9	34
6	M	32+3	41+4	2.50	1.90	Hirschprung's disease confirmed by rectal biopsy Struggling to open bowels with washouts	10	10	22.8	31
7	F	37+6	41+1	3.04	2.67	Gastroschisis Silo closure 20 days prior to imaging	180	10	25.3	30
8	F	32+1	39+1	2.65	1.63	Correction of type 2 duodenal atresia/stenosis by oesophagogastroduodenoscopy 47 days prior to imaging	90	20	24.9	32
10	M	25+3	44	2.50	0.740	NEC with ileal microperforation Laparotomy and ileostomy formation 108 days prior to recording Short gut syndrome	240	10	24.4	36

Acknowledgement