

# Personalized risk prediction for breast cancer pre-screening using artificial intelligence and thermal radiomics

Citation for published version (APA):

Kakileti, S. T., Madhu, H. J., Manjunath, G., Wee, L., Dekker, A., & Sampangi, S. (2020). Personalized risk prediction for breast cancer pre-screening using artificial intelligence and thermal radiomics. *Artificial Intelligence in Medicine*, 105, Article 101854. <https://doi.org/10.1016/j.artmed.2020.101854>

**Document status and date:**

Published: 01/05/2020

**DOI:**

[10.1016/j.artmed.2020.101854](https://doi.org/10.1016/j.artmed.2020.101854)

**Document Version:**

Publisher's PDF, also known as Version of record

**Document license:**

Taverne

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

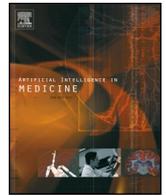
[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.



## Personalized risk prediction for breast cancer pre-screening using artificial intelligence and thermal radiomics<sup>☆</sup>

Siva Teja Kakileti<sup>a,b,\*</sup>, Himanshu J. Madhu<sup>a</sup>, Geetha Manjunath<sup>a</sup>, Leonard Wee<sup>b</sup>, Andre Dekker<sup>b</sup>, Sudhakar Sampangi<sup>c</sup>

<sup>a</sup> Niramai Health Analytix Pvt Ltd., Koramangala, Bangalore, Karnataka, India

<sup>b</sup> Department of Radiation Oncology (MAASTRO Clinic), GROW School for Oncology and Developmental Biology, Maastricht University Medical Centre +, Maastricht, The Netherlands

<sup>c</sup> Department of Radiology, Health Care Global, Bangalore, India

### ARTICLE INFO

#### Keywords:

Breast cancer  
Thermography  
Risk assessment  
Machine learning  
Thermalyx  
Artificial intelligence

### ABSTRACT

**Motivation:** Breast cancer is the leading cause of cancer deaths among women today. Survival rates in developing countries are around 50%–60% due to late detection. A personalized, accurate risk scoring method can help in targeting the right population for follow-up tests and enables early detection of breast abnormalities. Most of the available risk assessment tools use generic and weakly correlated features like age, weight, height etc. While a personalized risk scoring from screening modalities such as mammography and ultrasound could be helpful, these tests are limited to very few metropolitan hospitals in developing countries due to high capital cost, operational expenses and interpretation expertise needed for a large screening population.

**Methods:** We propose and analyze a new personalized risk framework called Thermalyx Risk Score (TRS) to identify a high-risk target population for regular screening and enable early stage breast cancer detection at scale. This technique uses Artificial Intelligence (AI) over thermal images to automatically generate a breast health risk score. This risk score is mainly derived from two sub-scores namely, vascular score and hotspot score. A hotspot score signifies the abnormality seen from irregular asymmetric heat patterns seen on the skin surface, whereas vascular score predicts the presence of asymmetric vascular activity. These scores are generated using machine learning algorithms over medically interpretable parameters that describes the metabolic activity inside the breast tissue and indicate the presence of a possible malignancy even in asymptomatic women.

**Results:** The proposed personalized risk score was tested on 769 subjects in four breast cancer screening facilities. The subjects' age ranged from 18 to 82 years with a median of around 45 years. Out of the 769 subjects, 185 subjects were diagnosed with a breast malignancy by an expert radiologist after mammography, ultrasound and/or histopathology. Our personalized AI based risk score achieved an area under the receiver-operator curve (AUC) of 0.89 when compared to an age normalized risk score that showed an AUC of 0.68. We also found that if the computed risk score is used to place individuals into four risk groups, the likelihood of malignancy also increases monotonically with the risk grouping level.

**Conclusion:** The proposed AI based personalized risk score uses breast thermal image patterns for risk computation and compares favorably to other generic risk estimation approaches. The proposed risk framework solution is automated, affordable, non-invasive, non-contact and radiation free and works for a wide age range of women from 18 to 82 years, including young women with dense breasts. The proposed score might be further used to assign subjects into one of the four risk groups and provide guidance on the periodicity of screening needed. In addition, the automatically annotated thermal images localizes the potential abnormal regions and might empower the physician to create a better personalized care.

<sup>☆</sup> This article belongs to Special issue: Artificial Intelligence in Breast Cancer Care.

\* Corresponding author at: Niramai Health Analytix Pvt Ltd., Koramangala, Bangalore, 560095 Karnataka, India.

E-mail address: [sivateja@niramai.com](mailto:sivateja@niramai.com) (S.T. Kakileti).

## 1. Introduction

The worldwide incidence of breast cancer has been increasing every year. It is the largest cause of cancer-related deaths among women. According to the World Health Organization, approximately 627,000 women lost their lives to this disease in 2018 [1,2]. In developing countries like India, there are about 163,000 new cases every year and the survival rate is merely 50 % [2]. To tackle this high mortality rate and to detect breast cancer earlier, screening facilities and public access programmes are being improved in these countries [3]. At present, baseline incidence and mortality of female breast cancer in the Indian population are approximately 25.8 and 12.7 [2] respectively, per 100,000 women. This incidence rises to 35–40 per 100,000 in metropolitan cities like Delhi, Chennai and Bangalore [4]. The problem is complicated by a critical imbalance of supply - 1 radiologist per 100,000 persons in India - indicating that a lack of access to expertise and screening facilities plays a major role in breast cancer mortality [3]. A risk profiling mechanism that could identify high-risk persons earlier may help to focus personalized screening recommendations and thus making public screening programmes more effective.

In the current literature, there are three prominent epidemiological risk models that estimate the risk of breast cancer based on demographics, clinical and family cancer information. The Gail model [5] considered age, age at menarche, age at first childbirth, race, ethnicity, number of first-degree relatives with history of breast cancer and the number of breast biopsy examinations. The Claus model [6] assessed hereditary risk of breast cancer by including comprehensive family cancer history information. The Tyrer-Cuzik model [7] was a recent risk estimation model that accounted for genetic mutations such as BRCA1 and BRCA2 genes, atypical hyperplasia, lobular carcinoma *in situ* and the other parameters in the Gail model.

Though the above risk models can be used for identifying a target population for screening, they are only moderately personalized and have a weak correlation to actual screening outcome [12]. As most of the features used in these models are generic, the models group participants into large cohorts as high risk or low risk, where the malignancy incidences in these cohorts might not significantly correlate with their risk. In addition, obtaining all risk factors needed for the model is challenging due to privacy issues, fear of social discrimination or just lack of data. With missing information, the risk score estimated by these models might not be accurate in grouping the women into high or low risk.

In this paper, we introduce a new methodology of personalized risk prediction for breast cancer pre-screening using breast thermal patterns. We discuss and analyze an automated scoring framework called Thermalytix Risk Scoring (TRS) that uses machine learning analysis on breast thermal imaging to estimate breast cancer risk. Our hypothesis is that since thermal imaging provides information about abnormal metabolic activities in the breast region [8], these images may be used to estimate breast cancer risk enabling personalized screening. Specifically, we hypothesize that TRS can identify a high-risk target population who can be recommended for regular screening to detect breast cancers in their initial stages.

## 2. Background

Thermography is an emerging modality that involves image analysis of heat patterns emitted by a human body. In the presence of a breast malignancy, the metabolic activity in the immediate vicinity increases due to high resource consumption by the tumor as well as due to increased blood flow from existing and newly formed blood vessels (neovascularization) [8,9]. This metabolic activity is facilitated by the release of nitric oxide that is autocatalyzed by ferritin ions [10] and the heat generated is transferred to the breast skin surface through venous convection and conduction. With the advanced infrared cameras that can detect temperature differences as low as 0.05 °C, it is now possible

to identify, analyze and differentiate changes in heat patterns due to metabolic and vascular abnormalities. Since these metabolic perturbations start from the onset of cancer cells long before they coalesce into a solid mass, these heat patterns may be used to detect a sub-clinical cancer in the initial stages. In a research study by Gautherie [11], it was pragmatically validated in a large population that thermal imaging could detect breast cancers up to 5 years earlier than conventional screening modalities.

Though conventional screening techniques like mammography can be used to create personalized care [12], it has lot of challenges to be adapted to developing countries. Mammography uses X-rays to image the physical tissue density variations in the breast region. Solid tumors have a high density and therefore appear as enhanced regions in mammograms. This explains its proven low accuracy in younger women, where the surrounding breast tissue has a higher density and thus image contrast is poorer. Mammography is painful as it involves the compression of breast with 10–20 pounds of pressure to obtain the necessary images. The use of x-radiation also limits the number of screenings that can be taken by women over their lifetime and especially at a young age, as each exposure increases the cumulative risk of radiation-induced cancer.

Thermal imaging can be used a limitless number of times and is a non-contact, painless screening method, since it only collects the non-ionizing thermal radiation emitted from the subject's skin surface. Its efficacy has been tested in the field for women of all ages, and heat conduction is even more efficient for younger women with dense breast tissue (e.g. younger women). Even with advanced thermal cameras, thermography investment costs are about 5,000–25,000 USD which is less than 1/10th of conventional mammography. Furthermore, the method is highly portable and does not require high voltage power supplies which may be hard to find in remote areas. Thermal cameras are compact and can be easily set up as "screening rooms" in the comfort and privacy of a domestic residence. In our implementation, it is possible to make the breast screening completely private by triggering the camera shutter remotely and using artificial intelligence to analyze the anonymous images in real time.

There are several studies that tested the efficacy of thermography for breast cancer. In 1982, the US Food and Drug Administration (FDA) approved thermography for breast cancer as an adjunct modality. The first breast cancer thermography trial dates to the 1960's, where Gershon-Cohen et al. [13] were able to identify malignancies from thermal images with a sensitivity and specificity of 91.6 % and 92.4 %, respectively, in 1924 subjects. Stark and Way [14] reported a sensitivity of 98.3 % and 93.5 % on 4621 subjects. Spitalier et al. [15] reported 89 % sensitivity and 89 % specificity on 61,000 women with a longer 10-year period follow-up. In this study, they also reported that thermography showed early signs in 60 % of the cancerous cases. In similar studies conducted by Haberman et al. [16] and by Gros and Gautherie et al. [11], they attained high sensitivities close to 90 % on 39,800 and 58,000 subjects by following up subjects for 3-year and 5-year periods, respectively. Both studies observed that approximately 30 % of cases found an initial abnormality in breast thermograms before any other traditional screening techniques. Lastly, in Rassiwal et al. [17] thermography was able to classify breast cancers accurately with a sensitivity and specificity of 97.6 % and 99.2 %, respectively, on 1008 women. However, the complication in all these studies was the manual expertise needed to interpret the breast thermograms. The analysis of breast thermal images is not trivial as the changes in the breast heat patterns can also be seen with benign conditions like fibroadenoma, duct ectasia etc. In addition, the breast thermal images were generally represented using false color palettes and the experts needed to pick out the thermal properties of malignancies visually from these false color images. Due to these reasons, results were highly dependent on the expertise of manual thermographers. This dependence on expert interpretation was one of the main reasons for poorly reproducible results of breast thermography in the Breast Cancer Demonstration Project

(BCDDP), which compared different screening techniques for breast cancer [8,18].

### 3. Role of machine learning in thermal imaging

Machine learning (ML) is the process of iteratively training an algorithm to produce a correct response to a stimulus (input) without explicitly having that response programmed into the algorithm from the outset. In the field of healthcare, ML has shown its efficacy in achieving high accuracies, scaling and reproducibility. Analyzing medical images typically involves the extraction of a variety of features that could reveal the disease characteristics (radiomics) followed by an ML classifier that predicts the disease outcome based on these features.

Thermal radiomics deals with the features that quantify breast thermal patterns and their associations with cancerous metabolism. These features could empower the physician in quantifying the disease nature and provide additional information that might not be detected with the unaided eye. The automation of cancerous detection with these features could assist with reproducible classification and reduce the sole dependence on human experts. It could also enable large scale screenings without the need for a vast army of human experts in every part of the country. Borchardt et al. [19] summarized the academic studies on the efficacy of machine learning algorithms for breast cancer thermal screening. These techniques involve the use of textural radiomic features such as energy, uniformity, correlation, entropy, among others, for differentiating the malignant patterns from normal heat distribution. The main drawback of these features is that they use the entire breast region for analysis and do not localize the abnormal regions spatially. Hence, these features describe the overall heat pattern of the breast and do not give specific information about the abnormality, something which could help the physician in understanding the nature of tumor.

Our own earlier research on using machine learning for breast thermography has shown promising accuracy results [20–23]. In [20], we described the effectiveness of medically interpretable imaging features obtained from the abnormal thermal patterns to get high levels of sensitivity and specificity of interpretation. In [21], we proposed the use of shape-based and temperature-based image processing features to detect the vasculatures from thermal images. These algorithms have shown comparable and sometimes better results than standard of care [22,21–23].

To summarize, there are very few risk assessment solutions to generate a personalized risk score for pre-screening of breast cancer. Most of these techniques use demographic and clinical features as discussed above. The use of mammography for personalized risk score is not ideally suited for developing countries due to its aforementioned challenges. In this paper, we discuss a novel solution called the Thermalyx risk scoring framework that generates a breast health risk

score based on the thermal images and further localizes the abnormal regions for radiologists to possibly support personalized care.

### 4. Methodology

In order to quantitatively give a risk score for breast cancer, we propose a risk assessment tool which quantifies heat patterns on the skin surface of the breast. We use machine learning algorithms to extract, detect and characterize radiomic features in the thermal images. We estimate a risk score called Thermalyx Risk Score (TRS) using two sub-scores: vascular score and hotspot score. A vascular score helps in detecting the abnormal properties in the region related to blood flow and neo-angiogenesis, while the hotspot score helps in describing the phenotype properties of a localized abnormality. In this sense, we provide a quantitative analysis of thermal images which signify physiological activities of the breast tissue. This analysis is more objective than human experts who mentally segregate the heat patterns into thermally locally active regions (hotspots) and adjacent vascularity features [8,18]. A description of the methods used for generation of these scores follows.

#### 4.1. Computing hotspot score

A malignant tissue generates more heat due to the release of nitric oxide in its immediate location, which is detected as a highly localized region of elevated temperature (i.e. hotspots) in the thermal images. However, nitric oxide is also released due to localized inflammation and other benign conditions such as fibro-adenoma, ductal ectasia, etc. The operating hypothesis of thermography is that the heat patterns from a breast malignancy must be differentiated from other conditions due to their differences in the biological origin [8]. To characterize these heat patterns, we extracted different image features that could help in differentiation between malignant and non-malignant conditions.

The first step in abnormal hotspot feature extraction is the detection of the hotspot regions. As shown in our earlier work [20], fusion of multiple temperature-based thresholds performs well for segmenting these abnormal regions. It is also noted that segmenting the regions that are less thermally active (warm spots) could help in analyzing the benign and deeper malignant lesions. Hence, we use  $T_1$  and  $T_2$  as shown below to extract the hotspot and warm spot regions by choosing different  $\rho$ ,  $\tau$ .

$$T_1 = \mu + \rho * (Tmax - \mu) \quad (1)$$

$$T_2 = Tmax - \tau \quad (2)$$

In the above equation,  $\mu$  and  $Tmax$  represents mean and maximum of the breast temperatures across all the views respectively. The parameters ( $\rho$ ,  $\tau$ ) are chosen such that they maximize the weighted average of sensitivity and specificity for malignancy classification [20]. We

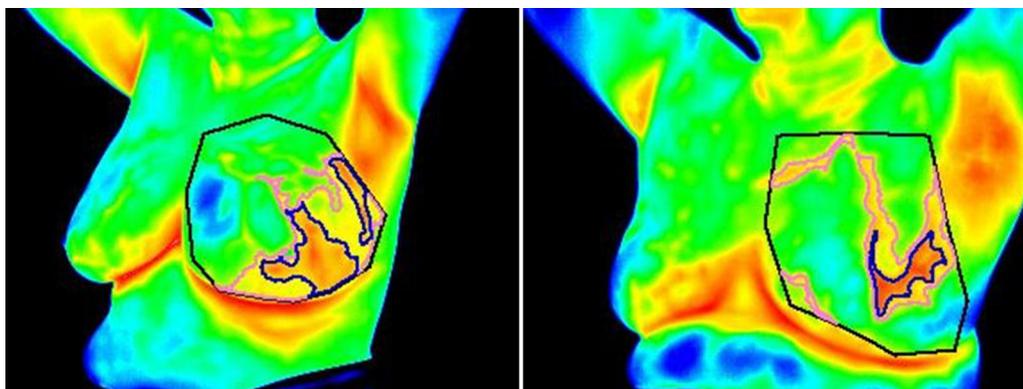


Fig. 1. Thermal image of two malignant subjects with hotspot and warmspot regions represented in blue and pink boundaries respectively (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

found that the  $(\rho, \tau)$  values of (0.56,1.9) and (0, 2.4) for hotspots and warmspots respectively results in maximum weighted average through grid search [20]. Fig. 1 shows annotated thermal images with detected hotspot and warm spot regions for two subjects with a malignancy. Once these regions are detected, medically relevant features are extracted to characterize these abnormal regions as detailed in [20]. These features can be categorized as below:

#### 4.1.1. Boundary and shape features

Benign lesions are confined to tissue boundaries, whereas the aggressiveness of cancerous lesions makes them rupture their tissue boundaries and invade the surrounding regions [8]. Hence, the use of features like deviation from natural shapes like circle/ellipse, irregularity and fractal dimensionality quantifies the structural properties of the lesion and help in differentiating the regions into benign and malignant lesions.

#### 4.1.2. Relative temperature

The temperatures of the detected abnormal regions correspond to their metabolic activities. Human thermographers consider any region with 2 °C–2.5 °C increase in the temperature as suspicious for malignancy [8,18]. To incorporate this domain knowledge, we considered relative increase of mean temperature of these abnormal regions with respect to mean temperatures of the surrounding region and contralateral breast regions as the features.

#### 4.1.3. Symmetry

Symmetry between breasts is a well-known feature in breast radiology. Symmetry can help in delineating the normal and abnormal changes. Any normal change has high possibility of occurring bilaterally. So, we have extracted features like mirror overlap, thermal distribution ratio and area difference as described in [20] to study the symmetry.

#### 4.1.4. Presence and extent of abnormal regions

This corresponds to the existence and the extent of hotspot and warm spot regions in the breasts.

Overall, 31 features were extracted from the detected abnormal regions as shown in Table 1. Hotspot features are extracted from the view having maximum hotspot region (best view) and whereas the warmspot features are extracted from both best view and its contralateral view. To differentiate the malignant heat patterns, we used a random forest (RF) classifier with 200 decision trees, which has shown better accuracies for breast thermography in our previous works [20]. RF is an ensemble of distinct decision trees whose outputs are pooled together to obtain the final classification. A hotspot score is then calculated as the fraction of trees out of the total number, that classify the features as malignant. The RF is initially trained on a labelled data set whose class labels (malignant or benign) are obtained using mammo-graphy, ultrasound and biopsy results.

## 4.2. Computing vascular score

Li et al. [24] observed that vascular changes happen in the very initial stages of cancer much before they coalesce into a solid mass. These vascular changes help in feeding the cancerous cell with the

required resources. Therefore, identifying the vascular changes could help in early detection of cancer. Since more volume of blood flows in these vessels with higher velocity, they generate more thermal activity and can be captured using thermal infrared cameras. Unlike many segmentation problems where there are distinct boundaries, thermal vessel signatures are diffused as they represent the transmitted heat from the vessels that reside beneath the breast skin surface. Hence, we convolve the thermal image with three different variations of gaussians for enhancing the vessel boundaries as discussed in our previous work [21]. Each of these enhanced images are then passed to a shape and temperature filter for identifying the pixels with high temperatures and which have a shape close to vessel like structures.

Shape filters detect the pixels that have a high curvature in one direction and low curvature in its perpendicular direction using Eqs. (3) and (4).  $V_s$  is calculated for different scales from  $s_{min}$  to  $s_{max}$  and the mode of these scales is taken to detect the pixels having vessel (tube) like shape.

$$V_s = \begin{cases} \text{sign} \left( \text{sinc} \left( \frac{|\lambda_{2s}|}{|\lambda_{1s}| - |\lambda_{2s}|} \right) - t \right) \lambda_{1s} < 0 \\ 0 \text{ otherwise} \end{cases} \quad (3)$$

$$V = \text{mode}_{s_{min} \leq s \leq s_{max}} V_s \quad (4)$$

In the above equation,  $\lambda_{1s}$  and  $\lambda_{2s}$  ( $\lambda_{1s} \geq \lambda_{2s}$ ) correspond to the maximum and minimum eigenvalues of a Hessian matrix at scale  $s$ , respectively, and  $t$  is a parameter that decides the threshold for vessel and non-vessel regions.  $V_s$  represents vessel-ness measure at scale  $s$  and  $s_{min}$  and  $s_{max}$  corresponds to the minimum and maximum possible thickness,  $s$ , of vessels.

On the other hand, a temperature high pass filter identifies the pixels that have higher temperatures compared to their surrounding mean temperature. Thus, an intersection of both these temperature and shape filters results in pixels that have both vessel structure and locally maximum temperatures. On obtaining these intersected outputs for each of the enhanced image, they are all combined to give the final vessel pixels [21].

Once the vessels are segmented, features quantifying the nature of vessels are extracted for further classification. These features include tortuosity, number of vessels, number of branches, extent of vessels, symmetry of vessels in both breasts and temperature increase of vessels w.r.t to the surrounding region and contralateral breast regions. In total, we extracted 21 features from these vessels detected in both the breast regions. These features are then fed to an RF classifier with 200 decision trees as discussed in previous subsection for binary classification to obtain the vascular risk score.

## 4.3. Computing the Thermalytix Risk Score

To obtain the final Thermalytix risk score (TRS), we have combined the proposed vascular score, hotspot score with critical symptoms like lump and nipple discharge as shown in below equation:

$$\text{TRS} = \begin{cases} 0.25 + \left( \frac{0.75}{2} \right) * (\text{hotspot score} + \text{vascular score}) & \text{if lump or nipple discharge} \\ 0.8 * \text{hotspot score} + 0.2 * \text{vascular score} & \text{else} \end{cases} \quad (5)$$

In case of symptoms such as lump or discharge, equal weight is

**Table 1**

The 31 features extracted from both hotspot and warmspot regions.

| Properties                    | Features Extracted  | Hotspot features count | Warmspot features count |
|-------------------------------|---|------------------------|-------------------------|
| Boundary Features             | Deviation from circle and ellipse, Irregularity and Fractal dimensionality. | 4                      | 8                       |
| Contralateral-side Comparison | Mirror overlap, Thermal distribution ratio and Area difference.             | 5                      | 5                       |
| Relative Temperature          | Relative temperature to surrounding tissues                                 | 1                      | 2                       |
| Presence of abnormal regions  | Number and Size   | 2                      | 4                       |

given to both the hotspot and vascular scores and an additional risk score of 0.25 is added to signify the risk of these symptoms. In the absence of lump, a lower weight is given to the vascular score in order to reduce the impact of high scores produced during natural hormonal activities. The 4:1 weightage ratio of hotspot score to vascular score is found from our previous experiments by performing a grid search of weights similar to hotspot parameters calculation.

### 5. Dataset description

To compare the performance of our proposed approach, we determined the TRS on 769 subjects who visited for either mammography or ultrasound screening at four clinical sites consisting of 3 cancer hospitals (581 women) and one diagnostic centre (188 women). Pregnant women, lactating mothers and cancer survivors were excluded from the data collection. We performed breast thermography imaging prior to any conventional diagnostic test for each of these subjects. All persons were informed about the study and enrolled after signing a consent form. Our breast thermography acquisition protocol involved removing the external clothing from the breasts and cooling the subject for 10–15 min to remove any extraneous heat. This was followed by thermal image capture at 5 different viewing angles, which took 2 min per person. Fig. 2 summarizes these different steps involved in the thermal image capture and TRS computation. The subjects then continued onto the other diagnostic tests as recommended by the attending radiologists at the respective data collection sites. The final ground truth is assumed to be the conclusion given by the attending radiologists after mammography, ultrasound and/or histopathology. Radiologists remained fully blinded to the thermography results for the duration of the study. Mammography was only available for 407 subjects due to the high tissue density in the remaining subjects or as per radiologist recommendation in case of a painful breast that was unsuitable for compression. Majority of the subjects underwent ultrasound imaging and we obtained at least one of either mammography or ultrasound reports for each of the enrolled subjects. Biopsy or Fine Needle Aspiration Cytology (FNAC) was recommended in case of suspicious findings in these reports. Out of 769 subjects, 185 subjects were determined to have a breast malignancy by the radiologist after looking at the available reports. 100 out of these 185 malignancies were histopathologically confirmed for malignancy either through FNAC or

**Table 2**  
The detailed age distribution of the subjects who participated in the study.

| Age   | Normal       |      |           | Malignant    |      |           |
|-------|--------------|------|-----------|--------------|------|-----------|
|       | No. of cases | Lump | Screening | No. of cases | Lump | Screening |
| 15–30 | 59           | 28   | 14        | 2            | 1    | 0         |
| 30–35 | 69           | 36   | 16        | 6            | 5    | 0         |
| 35–40 | 82           | 36   | 23        | 23           | 19   | 0         |
| 40–45 | 113          | 38   | 48        | 18           | 16   | 0         |
| 45–50 | 82           | 15   | 39        | 33           | 29   | 1         |
| 50–55 | 75           | 26   | 32        | 22           | 16   | 1         |
| 55–60 | 49           | 7    | 25        | 31           | 22   | 5         |
| 60–65 | 29           | 4    | 17        | 18           | 16   | 0         |
| > =65 | 26           | 8    | 10        | 32           | 26   | 4         |

Biopsy in addition to the imaging tests. This high number of malignancies is due to the selection of 3 cancer hospitals for this study.

The detailed distribution of the subjects is given in Table 2. Overall, 534 subjects came with complaints such as nipple discharge, pain, lump etc. and the remaining 235 subjects had no complaints and came for regular breast screening.

### 6. Results

The first step is to find the cropped breast region automatically from the thermal image. For this, an initial prediction was estimated using a V-net architecture as discussed in our earlier work [25] and then refined manually by human experts. For all the 769 subjects we calculated the hotspot and vascular scores as discussed in the previous sections and our earlier work [20,21]. The  $t$  value is set to 0.1 to identify the pixels with vessel shape and we have considered the mode over 0 ( $S_{min}$ ) to 10 ( $S_{max}$ ) scales as described in our prior work [21]. The RF classifiers were pretrained using a prior collected dataset consisting of 128 subjects with 51 malignancies. In this study, we used these pretrained classifiers and the entire 769 subjects were used as test set for validating the TRS.

We present the results in two-fold: (a) Effectiveness of the TRS using ROC analysis by comparing with age normalized risk score (b) Comparison of the TRS with the actual malignancy status of the subjects.

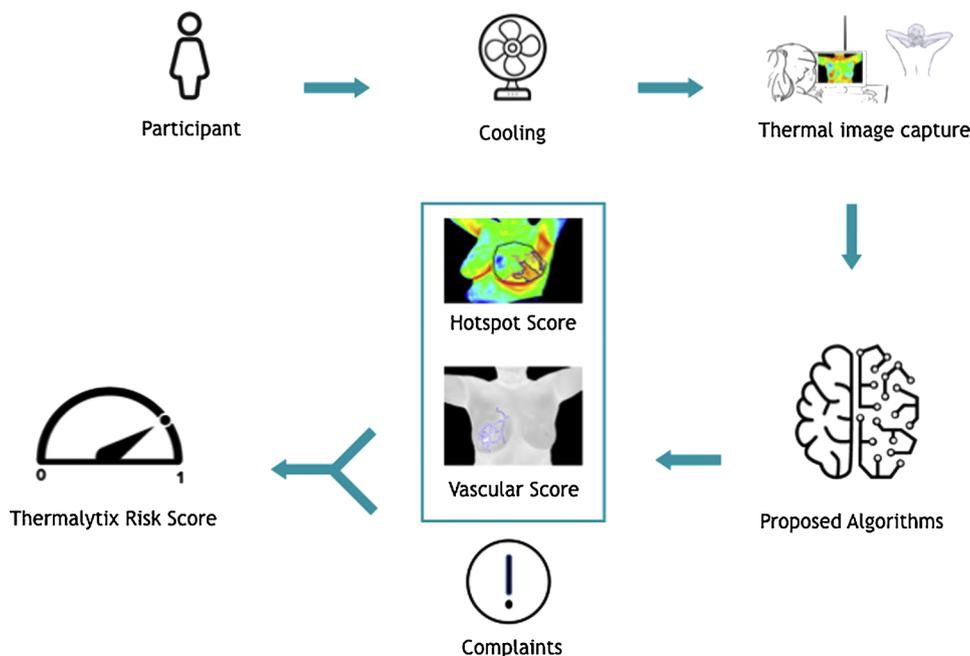


Fig. 2. The overall process involved in generating final Thermalytix Risk Score.

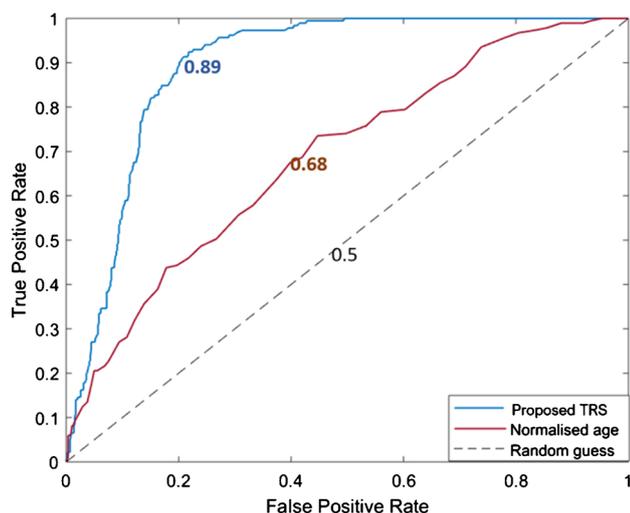


Fig. 3. Receiver Operating Characteristic (ROC) curve for age normalized risk and the proposed TRS.

6.1. Evaluating the effectiveness of TRS

The proposed TRS resulted in an Area Under receiver operating characteristic Curve (AUC) of 0.895 as shown in Fig. 3. This AUC is superior to the age normalized score that produced an AUC of 0.68. Age normalized score is computed by normalizing the age with minimum and maximum age values in the obtained dataset. Age is the most common criterion used to assess the risk in the developing and under-developed countries due to the lack of risk assessment tools tuned to these countries.

The dataset did not have any information about the BRCA genes, detailed family history, race, age at menarche, body mass index (BMI) etc. Due to this, the models like Gail [5], Claus [6] and Tyrer-Cuzik [7] were not applicable in the current scenario.

6.2. Risk stratification using TRS

We stratified the study population into 4 uniform width risk cohorts based on their TRS as [0 – 0.25: 1, 0.25 – 0.5: 2, 0.5 – 0.75: 3, 0.75–1.0: 4]. Fig. 4a shows the number of subjects per risk level and the percentage of malignant subjects in each level. As seen, when we use 4 risk cohorts (risk cohort 1, 2, 3, 4), the likelihood of malignancy correlates with the risk level. No subject in the risk cohort 1 is found to be malignant. Approximately 12 % of the women in the risk cohort 2 were found to be malignant as per radiologist conclusion. This malignancy incidence has increased to 57 % and 65 % for the risk cohorts 3 and 4,

respectively. This shows that TRS is an effective method of risk scoring where higher TRS corresponds to higher risk of breast cancer.

Typically, women under 45 years of age are not recommended for regular screening by mammography due to low breast cancer incidences in this age group and low sensitivity of mammography in dense breasts. However, with the recent increase in the breast cancer incidence among younger women, one might argue for regular screening for this population as well. We found that 48 % of walk-ins at the data collection sites were of age below 45 years and 13 % of these subjects were found to have malignant breast cancer. In order to analyze the effectiveness of the TRS in young women, Fig. 4b shows the same risk stratification for 372 women with age less than 45 years. As seen, the proposed risk scoring identified 104 of these young women as high risk and very high risk, upon which a breast Magnetic Resonance Imaging (MRI) or ultrasound could have been recommended as follow-up.

7. Discussion

The main motivation of this paper is to propose a risk assessment framework that could detect a population at high risk of malignant breast cancer. As shown in the results section, the proposed risk score showed an AUC of 0.89 when compared to an age-based risk estimation that gives an AUC of 0.68. This shows the potential of the proposed risk framework in accurately predicting the breast cancer risk from breast thermal images. The use of sole breast thermal images for TRS computation also avoids the missing data problem that is seen with the other risk assessment tools [5–7].

Fig. 4 shows the effectiveness of TRS in stratifying the screening population into four different risk levels. These four risk levels might be further used to create a personalized screening regime as follows:

- $TRS < 0.25$  (low risk): This group corresponds to the subjects whose heat patterns look normal and do not have critical symptoms like lump or discharge. This explains the risk weight of 0.25 in TRS calculation so that these critical symptoms are not ignored during risk group assignment. In the current study, none of the subjects are found to have a malignancy by the radiologist when the proposed risk score is less than 0.25. Hence, an annual follow up with proposed technique could be recommended as a non-radiation-based pre-screening for these low risk women.
- $TRS < 0.5$  and  $TRS \geq 0.25$  (Moderate risk): Approximately 12 % of the subjects whose risk score is in this range are found to have a malignant cancer in the current study. Hence, an annual follow up with proposed approach could be recommended adjunct to ultrasound or mammography for these subjects.
- $TRS < 0.75$  and  $TRS \geq 0.5$  (High risk): In the current study, 57 %

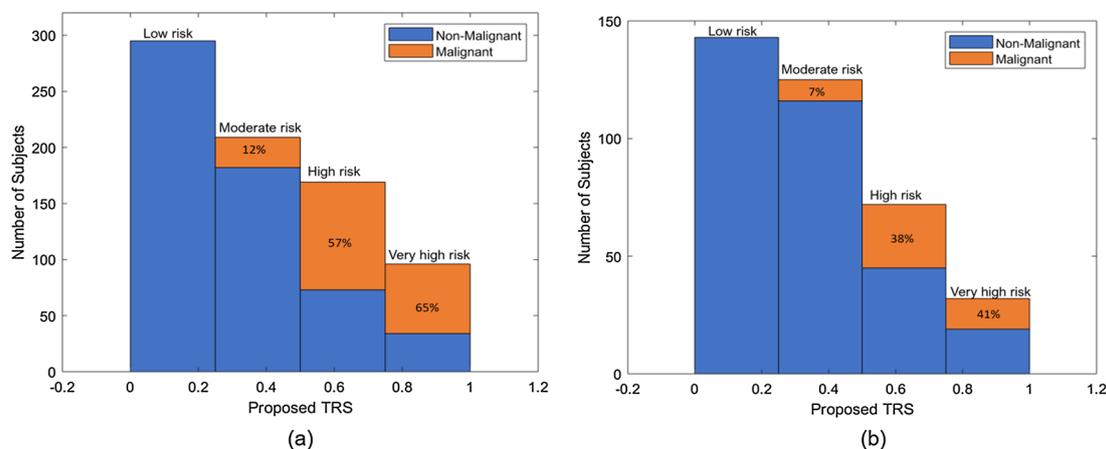


Fig. 4. The stacked bar plot of the risk score along with the total number of malignancies for (a) all 769 women (b) 372 women with under 45 years.

of the subjects falling in this range are found to have malignant cancer by the radiologist. Due to this high incidence, these subjects could be recommended for immediate follow up with other techniques like ultrasound/MRI/Mammography. In the case of negative results, an annual follow up with the proposed approach adjunct to ultrasound or mammography might be recommended for these subjects.

- $TRS \geq 0.75$  (Very High risk): These are high risk women with more than 65 % cancer incidence. An immediate follow up is recommended for these subjects especially at the localized abnormal regions. The immediate follow up can be with mammography or ultrasound or MRI or histopathology at detected abnormal regions. In the case of negative results, a short interval follow-up of 6 months with the proposed non-radiation approach might be helpful in detecting the malignancies at early stages [26].

The screening recommendations discussed above could help women to understand their risk and in identification of early stage breast cancers. Though approximately 40 % of women in high and very high risk categories are found to be negative by current standard of care, we believe these women are at high risk of breast cancer in future. Hence, we believe that the proposed screening regime for these high risk categories can help in detecting the cancers at early stages, thereby improving the chances of survival.

The high AUC with the proposed TRS also allows its usage as a pre-screening for breast cancer detection. The proposed TRS also provides the spatial location of abnormal regions along with their thermal radiomics as shown in Fig. 1. This spatial localization of the abnormal regions can help the physicians in creating a personalized diagnosis and treatment. It might guide the physician to look for presence of any physical abnormalities like lump, pain skin changes etc. at the detected regions using clinical breast examination. It might also help ultrasonographers in reducing the procedure time and might improve their performance by guiding them to the appropriate locations instead of recommending whole breast ultrasound imaging for all women. The spatial localization would be crucial especially for high and very high-risk cohorts, where Fine Needle Aspiration Cytology (FNAC) or histopathology can be done directly on these regions in case of suspicious findings with other modalities.

The current study validated the performance of Thermalytix on a small population of 769 women who visited our clinical sites comprising of three cancer hospitals and a breast cancer diagnostic centre in Karnataka, India. This choice of clinical sites led to 70 % symptomatic and 30 % asymptomatic women. It is also important to note that the 30 % asymptomatic women could be far from the real-world screening population as these were the women who walked into cancer hospitals and diagnostic centres. Therefore, to bring Thermalytix into the medical practice, we need to validate its performance on both screening and symptomatic populations with varying races, ages to remove any biases by conducting large scale studies.

## 8. Conclusion

In this paper, we have discussed a new risk assessment tool called Thermalytix Risk Score to estimate a personalized risk score that can be used as a pre-screening for breast cancer. The use of non-invasive, affordable and portable thermal imaging for predicting the risk score makes it more suitable for developing countries like India and may overcome some of the critical challenges faced by other breast cancer screening modalities. It works for a wide age range of women from 18 to 82 years including younger women with high breast tissue densities. The proposed TRS uses breast thermal patterns for risk estimation, thereby making it more personalized than generic risk estimation techniques e.g. based on age. The obtained results show that the TRS can estimate the breast cancer risk with a high area under curve compared to age-based risk score. The four risk categories discussed in this

paper can serve as the guidelines for creating screening procedures for detecting the malignancies in early stage. The proposed approach further detects the abnormalities in vascular structures and localizes the abnormal lesions in the thermal images. This might empower the doctor with better understanding of metabolic activity of different tissues inside the breast region and might direct more personalized care. This is an initial study and in future we intend to test the TRS on a larger and more diverse population.

## Research support

STK, HJM and GM are full-time employees of the private company Niramai Health Analytix where this research was done. Co-authors AD and LW receive academic research funding from an Indo-Dutch Joint Research Programme grant (no. 629.002.204) and a Dutch Technology Foundation research grant (no. 14930).

## Transparency document

The Transparency document associated with this article can be found in the online version.

## Declaration of Competing Interest

STK, HJM and GM have direct conflict of interest. AD, LW and SS have no direct or indirect conflict of interest.

## Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.artmed.2020.101854>.

## References

- [1] <https://www.who.int/cancer/prevention/diagnosis-screening/breast-cancer/en/>.
- [2] Bray Freddie, Ferlay Jacques, Soerjomataram Isabelle, Siegel Rebecca L, Torre Lindsey A, Jemal Ahmedin. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68(6):394–424.
- [3] Kalyanpur A. Commentary - radiology in India: the next decade. *Indian J Radiol Imaging* 2008;18:191–2.
- [4] Shreshtha Malvia, Bagadi Sarangadhara Appalaraju, Dubey Uma S, Saxena Sunita. Epidemiology of breast cancer in Indian women. *Asia Pacific J Clin Oncol* 2017;13(4):289–95.
- [5] Gail Mitchell H, Brinton Louise A, Byar David P, Corle Donald K, Green Sylvan B, Schairer Catherine, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *JNCI* 1989;81(24):1879–86.
- [6] Claus Elisabeth B, Risch N, Douglas Thompson W. Genetic analysis of breast cancer in the cancer and steroid hormone study. *Am J Hum Genet* 1991;48(2):232.
- [7] Amir Ethan, Evans DG, Shenton Andrew, Laloo Fiona, Moran Anthony, Boggis C, et al. Evaluation of breast cancer risk assessment packages in the family history evaluation and screening programme. *J Med Genet* 2003;40(11):807–14.
- [8] Kennedy Deborah A, Lee Tanya, Seely Dugald. A comparative review of thermography as a breast cancer screening technique. *Integr Cancer Ther* 2009;8(1):9–16.
- [9] Folkman Judah. What is the evidence that tumors are angiogenesis dependent? *JNCI* 1990;82(1):4–7.
- [10] Anbar Michael. Hyperthermia of the cancerous breast: analysis of mechanism. *Cancer Lett* 1994;84(1):23–9.
- [11] Gautherie M, Gros CM. Breast thermography and cancer risk prediction. *Cancer* 1980;45(1):51–6.
- [12] Yala Adam, Lehman Constance, Schuster Tal, Portnoi Tally, Barzilay Regina. A deep learning mammography-based model for improved breast cancer risk prediction. *Radiology* 2019:182716.
- [13] Gershon-Cohen J, Haberman-Brueschke JA, Brueschke EE. Medical thermography: a summary of current status. *Radiol Clin North Am* 1965;3(3):403.
- [14] Stark Agnes M, Way S. The screening of well women for the early detection of breast cancer using clinical examination with thermography and mammography. *Cancer* 1974;33(6):1671–9.
- [15] Spitalier H, Giraud D, et al. Does infrared thermography truly have a role in present day breast cancer management? In: Liss AR, editor. *Biomedical thermology*. 1982. p. 269–78. New York.
- [16] Haberman Jo Ann D, Love Francis TJ, John E. Screening a rural population for breast cancer using thermography and physical examination techniques: methods and results- A preliminary report. *Ann N Y Acad Sci* 1980;335(1):492–500.

- [17] Rassiwala M, Mathur P, Mathur R, Farid K, Shukla S, Gupta PK, et al. Evaluation of digital infra-red thermal imaging as an adjunctive screening method for breast carcinoma: a pilot study. *Int J Surg* 2014;12(12):1439–43.
- [18] Keyserlingk JR, Ahlgren PD, Yu E, Belliveau N, Yassa M. Functional infrared imaging of the breast. *IEEE Eng Med Biol Mag* 2000;19(3):30–41.
- [19] Borchardt TB, Conci A, Lima RCF, Resmini R, Sanchez A. Breast thermography from an image processing viewpoint: a survey. *Signal Processing* 2010;93(10):2785–803.
- [20] Himanshu Madhu, Kakileti Siva Teja, Venkataramani Krithika, Jabbireddy Susmija. Extraction of medically interpretable features for classification of malignancy in breast thermography. 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2016:1062–5.
- [21] Kakileti Siva Teja, Venkataramani Krithika. Automated blood vessel extraction in two-dimensional breast thermography. 2016 IEEE International Conference on Image Processing (ICIP) 2016:380–4.
- [22] Sudhakar S, Geetha Manjunath, Teja Kakileti Siva, Himanshu Madhu. Thermalytix: an advanced artificial intelligence based solution for non-contact breast screening. *Int J Med Health Sci Res* 2018;12(2):48–51. WASET.
- [23] Manjunath G, Sudhakar S, Kakileti S, Madhu H, Singh A. Abstract P6-02-12: artificial Intelligence over thermal images for radiation-free breast cancer screening. *Cancer Res* 2019. P6-02.
- [24] Li Chuan-Yuan, Shan Siqing, Huang Qian, Braun Rod D, Lanzen Jennifer, Hu Kang, et al. Initial stages of tumor cell-induced angiogenesis: evaluation via skin window chambers in rodent models. *J Natl Cancer Inst* 2000;92(2):143–7.
- [25] Kakileti Siva Teja, Manjunath Geetha, Madhu Himanshu J. Cascaded architecture for View independent breast segmentation in thermal Images. 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2019:6294–7.
- [26] Bowles Erin J Aiello, Miglioretti Diana L, Sickles Edward A, Abraham Linn, Carney Patricia A, Yankaskas Bonnie C, et al. Accuracy of short-interval follow-up mammograms by patient and radiologist characteristics. *Am J Roentgenol* 2008;190(5):1200–8.