Quantitative thermal imaging biomarkers to detect acute skin toxicity from breast radiation therapy using supervised machine learning

SAEDNIA, Khadijeh, TABBARAH, Sami, LAGREE, Andrew, WU, Tina, KLEIN, Jonathan, GARCIA, Eduardo, HALL, Michael, CHOW, Edward, RAKOVITCH, Eileen, CHILDS, Charmaine <http://orcid.org/0000-0002-1558-5633>, SADEGHI-NAINI, Ali and TRAN, William

Published version

SAEDNIA, Khadijeh, TABBARAH, Sami, LAGREE, Andrew, WU, Tina, KLEIN, Jonathan, GARCIA, Eduardo, HALL, Michael, CHOW, Edward, RAKOVITCH, Eileen, CHILDS, Charmaine, SADEGHI-NAINI, Ali and TRAN, William (2020). Quantitative thermal imaging biomarkers to detect acute skin toxicity from breast radiation therapy using supervised machine learning. International Journal of Radiation Oncology*Biology*Physics, 106 (5), 1071-1083.

Supplementary Table 1

Contains a list of the evaluated first order features along with their respective equations and descriptions.

First Order Feature	Equation	Description
Minimum Temperature	N/A	Minimum temperature value within the selected ROI in a pixel to pixel basis.
Maximum Temperature	N/A	Maximum temperature value within the selected ROI in a pixel to pixel basis.
Mean Temperature	Mean = $\frac{Sum\ of\ All\ Values}{Total\ Number\ of\ Values}$	Averaged temperature value
Median	N/A	The middle value from a set of ordered temperature values acquired from the selected ROI
Mode	N/A	The most reoccurring temperature value within the selected ROI
Entropy	S = $-\sum_i P_i Log P_i$	It measures the level of uncertainty of our system.
Skewness	$g_1 = \frac{\sum_{i=1}^{N}(Y_1 - \bar{Y})^3/N}{S^3}$	Is the measure of symmetry of the data set (ie. The distribution of temperature values within the ROI)
Kurtosis	Kurtosis = $\frac{\sum_{i=1}^{N}((Y_1 - \bar{Y})^4/N}{S^4}$	It measures whether the data are heavy tailed or light tailed.

P: probability; s: Standard deviation; Ȳ: Mean; N: Number of data points.

Supplementary Table 2

Contains a list of the evaluated GLRLM features along with their respective equations and descriptions (adapted from pyradiomics).

GLRLM Feature	Equation	Description	GLRLM Feature	Equation	Description
Short Run Emphasis (SRE)	$\frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\frac{\mathbf{P}(i,j\|\theta)}{j^2}}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\mathbf{P}(i,j\|\theta)}$	Measures short run distributions of homogenous pixels.	Run Variance (RV)	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}p(i,j\|\theta)(j-\mu)^2$ Where, $\mu=\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}p(i,j\|\theta)j$	Measures run length of pixels variance.
Long Run Emphasis (LRE)	$\frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\mathbf{P}(i,j\|\theta)j^2}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\mathbf{P}(i,j\|\theta)}$	Measures long run distributions of homogenous pixels.	Run Entropy (RE)	$-\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\frac{p(i,j\|\theta)}{\log_2(p(i,j\|\theta)+\epsilon)}$ Where, $\epsilon = A\ small\ possitive\ number$	Measures run length randomness and distribution of grey levels.
Grey level Non-Uniformity (GLN)	$\frac{\sum_{i=1}^{N_g}\left(\sum_{j=1}^{N_r}\mathbf{P}(i,j\|\theta)\right)^2}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\mathbf{P}(i,j\|\theta)}$	Measures grey values similarities between the pixels within an image.	Low Grey Level Run Emphasis (LGLRE)	$\frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\frac{\mathbf{P}(i,j\|\theta)}{i^2}}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\mathbf{P}(i,j\|\theta)}$	Measures low grey level value distribution within an image.
Grey level Non-Uniformity Normalized (GLNN)	$\frac{\sum_{i=1}^{N_g}\left(\sum_{j=1}^{N_r}\mathbf{P}(i,j\|\theta)\right)^2}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\mathbf{P}(i,j\|\theta)^2}$	Normalized GLN.	High Gray Level Run Emphasis (HGLRE)	$\frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\mathbf{P}(i,j\|\theta)i^2}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\mathbf{P}(i,j\|\theta)}$	Measures higher grey level value distribution within an image.
Run Length Non-Uniformity (RLN)	$\frac{\sum_{j=1}^{N_r}\left(\sum_{i=1}^{N_g}\mathbf{P}(i,j\|\theta)\right)^2}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\mathbf{P}(i,j\|\theta)}$	Measures run length similarities between the pixels within an image.	Short Run Low Grey Level Emphasis (SRLGLE)	$\frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\frac{\mathbf{P}(i,j\|\theta)}{i^2j^2}}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\mathbf{P}(i,j\|\theta)}$	Measures short run length joint distribution of pixels, with lower grey levels.

N_g: be the number of discrete intensity values in the image; N_r: be the number of discreet run lengths in the image; $\mathbf{P}(i,j|\theta)$: be the run length matrix for an arbitrary direction; $p(i,j|\theta)=\frac{\mathrm{P}(i,j|\theta)}{N_r(\theta)}$, normalized run length matrix; N_r: be the number of discreet run lengths in the image; θ: Angle.

Supplementary Table 2 (Cont.)

Contains a list of the evaluated GLRLM features along with their respective equations and descriptions (adapted from pyradiomics).

GLRLM Feature	Equation	Description	GLRLM Feature	Equation	Description
Run Length Non-Uniformity Normalized (RLNN)	$\frac{\sum_{j=1}^{N_r}\left(\sum_{i=1}^{N_g}\mathbf{P}(i,j\|\theta)\right)^2}{\left(\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\mathbf{P}(i,j\|\theta)\right)^2}$	Normalized RLNN.	Short Run High Grey Level Emphasis (SRHGLE)	$\frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\frac{\mathbf{P}(i,j\|\theta)i^2}{j^2}}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\mathbf{P}(i,j\|\theta)}$	Measures short run length joint distribution of pixels, with higher grey levels.
Run Percentage (RP)	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}\frac{\mathbf{P}(i,j\|\theta)}{N_p}$	Measures the distribution of homogenic pixels within an image.	Long Run Low Grey Level Emphasis (LRLGLE)	$\frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\frac{\mathbf{P}(i,j\|\theta)j^2}{i^2}}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\mathbf{P}(i,j\|\theta)}$	Measures long run length joint distribution of pixels, with lower grey levels.
Grey level Variance (GLV)	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}p(i,j\|\theta)(i-\mu)^2$ Where, $\mu=\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}p(i,j\|\theta)i$	Measures grey level intensity variances for the run length of pixels.	Long Run High Grey Level Emphasis (LRHGLE)	$\frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\mathbf{P}(i,j\|\theta)i^2j^2}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\mathbf{P}(i,j\|\theta)}$	Measures long run length joint distribution of pixels, with higher grey levels.

N_g: be the number of discrete intensity values in the image; N_r: be the number of discreet run lengths in the image; $\mathbf{P}(i,j|\theta)$: be the run length matrix for an arbitrary direction; $p(i,j|\theta)=\frac{\mathbf{P}(i,j|\theta)}{N_r(\theta)}$, normalized run length matrix; N_r: be the number of discreet run lengths in the image; θ: Angle; N_p: The number of pixels within the ROI.

Supplementary Table 3

Contains a list of the evaluated GLCM features along with their respective equations and descriptions (adapted from MATLAB:

GLCM Feature	Equation	Description
Energy (Ene)	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(p(i,j))^2$	Measures an image's homogenic patterns
Homogeneity (Hom)	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}\frac{p(i,j)}{1+(i-j)^2}$	Measures the closeness of the distribution of the primitives/elements in the GLCM to the GLCM diagonal.

graycoprops(glcm, properties)).

Contrast (Con)	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(i-j)^2 p(i,j)$	Measures the local grey tone intensity variations, prioritizing values away from the diagonal (i=j)
Correlation (Cor)	$\frac{\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(i,j)ij-\mu_x\mu_y}{\sigma_x(i)\sigma_y(j)}$	Measures neighbouring grey tone intensities linear dependency.

N_g: The image's number of discrete intensity values; $\mathbf{P}(i,j)$: Co-occurrence matrix for an arbitrary distance and angle; $p(i,j)=\frac{\mathbf{P}(i,j)}{\sum \mathbf{P}(i,j)}$, normalized co-occurrence matrix; $p_x(i)=\sum_{j=1}^{N_g}P(i,j)$, marginal rows probability; $p_y(j)=\sum_{i=1}^{N_g}P(i,j)$, marginal column probability; $\mu_x=\sum_{i=1}^{N_g}p_x(i)i$, mean grey level intensity of p_x; $\mu_y=\sum_{j=1}^{N_g}p_y(j)j$, mean gray level intensity of p_y; σ_x: Standard deviation of p_x; σ_y: Standard deviation of p_y.

Supplementary Table 4

Contains a list of the evaluated GTDM features along with their respective equations and descriptions (adapted from pyradiomics).

GTDM Feature	Equation	Description
Coarseness	$\frac{1}{\sum_{i=1}^{N_g} p_i s_i}$	Measures the average difference between the center pixel and it's surrounding neighborhood to discern spatial rate of change.
Contrast	$\left(\frac{1}{N_{g,p}(N_{g,p}-1)}\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p_i p_j (i-j)^2\right)\left(\frac{1}{N_{v,p}}\sum_{i=1}^{N_g} S_i\right)$ Where, $p_i \neq 0, p_j \neq 0$	Measures spatial intensity change and is dependant on overall grey level dynamic range. Contrast is high when the spatial intensity change rate and dynamic range are high.
Busyness	$\frac{\sum_{i=1}^{N_g} p_i s_i}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \lvert ip_i - jp_j \rvert}$ Where, $p_i \neq 0, p_j \neq 0$	Measures the changes that occur from a pixel to its neighbour.
Complexity	$\frac{1}{N_{v,p}}\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \lvert i-j \rvert \frac{p_i s_i + p_j s_j}{p_i + p_j}$ Where, $p_i \neq 0, p_j \neq 0$	An image can be denoted as complex when the image is non-uniform and there are many rapid changes in grey level intensity.
Strength	$\frac{\sum_{i=1}^{N_g}\sum_{J=1}^{N_g}(p_i + p_j)(i-j)^2}{\sum_{i=1}^{N_g} S_i}$ Where, $p_i \neq 0, p_j \neq 0$	A measure of the of the texture primitive/elements of an image. The more defined and visible these primitive/elements are, the higher the strength value.

$p_i = \frac{n_i}{N_v}$, grey level probability; n_i: The number of sets of segmented pixels pertaining to grey level i; N_v: The number of pixels within the ROI; N_p: The total number of pixels within the ROI; $N_{v,p}$ = The total number of sets of segmented pixels that are equal to $\sum n_i$, where $N_{v,p} \leq N_p$; N_g: The image's number of discrete intensity values; $N_{g,p}$: Number of grey levels ($p_i \neq 0$); $S_i = \begin{cases} \sum^{n_i} \lvert i - \bar{A}_i \rvert \text{ for } n_i \neq 0 \\ 0 \text{ for } n_i = 0 \end{cases}$, the summation of grey level i absolute differences; A: Surface area of mesh (mm^2).

Supplementary References

- Van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, Beets-Ta, RGH, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Research* 2017;77(21) doi: 10.1158/0008-5472.CAN-17-0339.