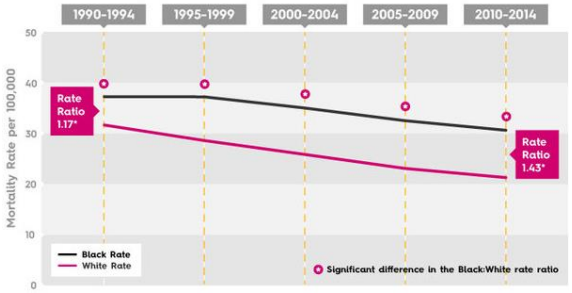
RIMMES RESEARCH FINAL REPORT

Author: Ranjini Ghosh Mentor: Dr. Jun Kong

INTRODUCTION

Clinically, the mortality rate for Black women diagnosed with breast cancer is 42% higher than the comparable rate for White women. In 1990, black women with breast cancer were 17% more likely to die than white women; nine cities presented statistically significant differences in mortality rates between black women and white women. By 2000, the disparity had risen to 35%; and 19 cities displayed statistically significant differences. Between 2010 and 2014, black women were 43% more likely to die than white women; and 24 cities presented statistically significant differences. Additionally, seventeen more cities showed the same trend.



OBJECTIVE

Understanding and rectifying these disparities are crucial steps toward achieving equitable healthcare outcomes for all women affected by breast cancer. This research project focuses on analyzing differences in the effects of breast cancer between black and white women by histology whole slide images (WSIs). The aim is to define and extract tumor microenvironment (TME) features from WSIs and mine for variations in the effect of breast cancer in these two races.

METHODS

The dataset consists of histological labels and corresponding spatial coordinates for various tissue types from Piedmont Hospital in Georgia.

Data Preprocessing: The dataset is filtered to include tissues histologically categorized as Adipose (1), Immune (2), Necrosis (3), Normal (4), Stroma (5), and Tumor (6). This refined dataset is converted into a Pandas DataFrame for ease of manipulation.

Delaunay Triangulation: The coordinates of tissue nodes are triangulated resulting in a Delaunay triangulation plot that provides a graphical representation of the spatial connectivity across different tissue types.

Triangle Area Calculation: The areas of the triangles formed by the Delaunay plot are computed. Such area information helps reveal spatial variations in distributions of tissues of different histology types.

Feature Extraction:

1. Class Distribution

This feature counts the occurrences (number of nodes) of each tissue type per total number of nodes. Each k x k sample of the same tissue type is defined as a node of that tissue type. In this way, such features can characterize tissue type distributions.

$$C(x,k,j) = \frac{n(x,k,j)}{N(k,j)}$$

C = Class distribution

x = tissue type (histology label)

k = side of neighborhood square

j = shifting pixel number

n = number of nodes of tissue type x

N = total number of nodes

2. Adjacency factor

This feature can be calculated for any two pairs of histology labels. It characterizes the spatial proximity of different tissue types.

$$A(k,m,n) = \frac{x(k,m,n)}{T(n)}$$

A = Adjacency factor

k = side of a neighborhood square

m = histology label for the center cell

n = histology label for cells in the neighborhood of m

x = number of tiles of label n around tiles of m

T = total number of tiles of label n

3. Density Feature

This feature represents the density of nodes from the Delaunay plot with triangle areas.

$$D(n) = \frac{no.\ of\ triangles\ with\ area\ less\ than\ n}{total\ no.\ of\ triangles}$$

Statistical Test (T-Test):

The t-test compares TME features between groups by race (Black and White) and socioeconomic status (Area Deprivation Index - ADI). It suggests that differences in ADI scores are statistically significant between groups. A low p-value (less than 0.05) suggests significant disparities.

Class Distribution Analysis:

Groups were created based on race (White or Black) and ADI brackets. Each race was further divided into four ADI brackets ([0-25), [25-50), [50-75), [75-100)).

Adjacency Factor Analysis:

Initial comparisons were made by race, then within each race across ADI brackets.

Density Feature Analysis:

Initial comparisons were made by race, then within each race across ADI brackets.

RESULTS

1. Class Distribution Analysis

Feature	Groups	T-statistic	P-value
C(3,1,1)	White (25-50) vs Black (25-50)	-2.3336078	0.03961632
C(5,1,1)	White (25-50) vs Black (25-50)	3.1195296	0.00975817
C(5,2,1)	White (25-50) vs Black (25-50)	3.1076403	0.0099673
C(5,3,1)	White (25-50) vs Black (25-50)	3.1083758	0.0099543
C(5,4,1)	White (25-50) vs Black (25-50)	3.091861	0.0102519
C(5,5,1)	White (25-50) vs Black (25-50)	3.08473919	0.0103831

Reject null hypothesis: There is a significant difference between the groups for Class Distribution C(3,1,1), C(5,1,1,), C(5,2,1), C(5,3,1), C(5,4,1) and (5,5,1).

2. Adjacency Factor Analysis

T-test results based on Race

Feature	T-statistic	P-Value
A(3,1,3)	2.010350395461	0.048612321480
A(3,1,4)	2.307055960838	0.024298222151

T-test results based on ADI Groups

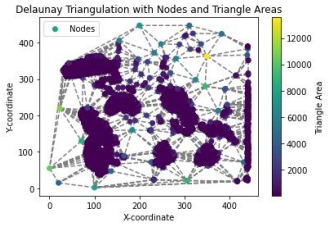
Feature	T-statistic	P-Value
A(3,1,3)	2.010350395461	0.048612321480
A(3,1,4)	2.307055960838	0.024298222151

Reject null hypothesis: There is a significant difference between the groups for Adjacency Factor A(3,1,3) and A(3,1,4)

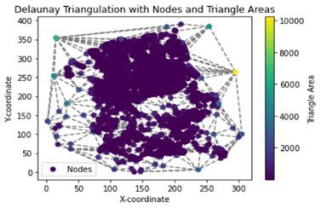
3. Density Feature Analysis:

Density Plots:

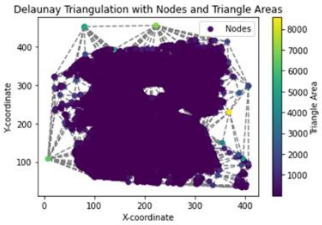
Race- White, ADI- 1



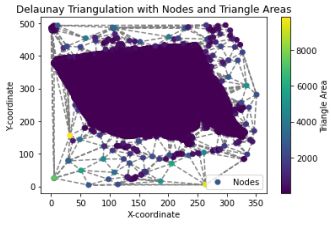
Race- Black, ADI- 1



Race- White, ADI- 88



Race- Black, ADI- 98



T-test results based on Race

Feature	T-statistic	P-Value
D(25)	-2.00380729	0.049327563

T-test results based on ADI

Feature	T-statistic	P-Value
D(25)	-2.00380729	0.049327563

Reject null hypothesis: There is a significant difference between the groups for D(25).

CONCLUSION

In summary, this research project aimed to explore disparities in the impact of breast cancer on black and white women through data science methodologies. By analyzing tumor microenvironment features extracted from patient histology WSI data, the study sought to identify differences in TME image features of breast cancer patients of two races and different ADI groups. Through the analysis of class distribution, adjacency factors, and density features, notable variations were observed. Addressing these variations is essential for advancing efforts toward equitable healthcare outcomes for all women with breast cancer.

REFERENCES

Justin Luningham, P Gaurav Seth, Geetanjali Saini, Shristi Bhattarai, Sofia Awan, Lindsay J. Collin, Monica Swahn, Dajun Dai, Keerthi Gogineni, Preeti Subhedar, Pooja Mishra, Ritu Aneja, Association of Race and Area Deprivation With Breast Cancer Survival Among Black and White Women in the State of Georgia

Babak Ehteshami Bejnordi, Jimmy Lin, Ben Glass, Maeve Mullooly, Gretchen L Gierach, Mark E Sherman, Nico Karssemeijer, Jeroen van der Laak, Andrew H Beck. Deep learning-based assessment of tumor-associated stroma for diagnosing breast cancer in histopathology images.

Harpinder Saini, Kiarash Rahmani, Jaimeson Veldhuizen, Azadeh Zare, Mayar Allam, Casey Silva, Alex Kratz, Danh Truong, Ghassan Mouneimne, Joshua LaBaer, Robert Ros, Mehdi Nikkhah. The Role of Tumor-Stroma Interactions on Desmoplasia and Tumorigenicity within a Microengineered 3D Platform.