The role of Age, Pathologic Stage, and Vital Status for Female Patients when Developing Breast Cancer

Mengdi Chai

05 December 2022

Introduction

Except for skin cancers, breast cancer is the most prevalent cancer in women in the US. 30% (or 1 in 3) of all new cases of female cancer each year are caused by it. Women in their middle years and older tend to develop breast cancer. At the time of breast cancer diagnosis, the average age is 62. As a result, 62 years of age or younger is the average age at which breast cancer in women is discovered. Women under 45 who have been diagnosed with breast cancer are extremely rare. In general, a woman's lifetime chance of acquiring breast cancer in the United States is around 13%. Because of the high prevalence and mortality rate of lung cancer, it is essential to discover the pathway and mechanism of breast cancer in female patients. The development of breast cancer is divided into several stages: stage I, stage II, stage III, stage IV, and stage X.

Breast cancer in stage I is an early stage and refers to a tiny, localized tumor that has only spread to the breast tissue or may have spread to nearby lymph nodes. In stage, the breat cancer can further be divided into IA stage and IB stage. A malignancy that is at stage IA is less than 2 centimeters (cm) in size and has not spread outside the breast, whereas Stage IB means that small areas of breast cancer cells are found in the lymph nodes close to the breast. Breast cancer that has reached stage II has spread to adjacent lymph nodes, the breast, or both. It is breast cancer in its early stages, and Stage II breast cancer has two groups: stage IIA and IIB. The classification of each group is depends on the size of tumour.

Stage III indicates that the cancer has migrated from the breast to nearby lymph nodes, the breast's surface, or the chest wall. Additionally, it is known as locally advanced breast cancer.Similarly, stage III also has different groups: stage IIIA, stage IIIB, and stage IIIC. Stage IIIA is quite complex since it depends on the size of tumour and the exist of breast cancer cells in the lymph nodes. Different from stage IIIA, Stage IIIB means the tumour has spread to the skin of the breast or the chest wall. In stage IIIC, there may or may not be a tumor, depending on its size. However, there is cancer in the skin of the breast that has gone to the chest wall and is causing edema or ulcers. Stage IV cancers have spread (metastasized) beyond the breast and nearby lymph nodes to other parts of the body.

The Cancer Genome Atlas Program (TCGA) is a cancer genomics program that collects over 20,000 cancer tissue samples and matches them with normal samples spanning over 33 cancer types (National Institute of Health). This publicly available database is used for multi-omic data analysis, which integrates data sets from many omic groupings.

In this study, I explored the survival rates of breast cancer in different age group at different pathologic stage to identify whether age and pathologic stage does play a role in patient outcomes. I hypothesized that age and pathologic stage will affect the female patient outcomes then they develop breast cancer.

Method

The analysis was mainly conducted in R using breast cancer clinical data. The data was sourced from TCGA with the TCGAbiolinks library using accession code "BRCA." In addition, another clinic

data was downloaded through TCGA website. Two data frames were merged together to access to "age_at_initial_pathologic_diagnosis", "ajcc_pathologic_stage", and "vital_status" columns.

In the dataset, there were a total of 2268 female patients. "age_category" column was created to divide female patients into young and old groups. According to the report of breast cancer, at the time of breast cancer diagnosis, the average age is 62. Therefore, female patients equal to or above 62 years old were categorized as "old", whereas female patients under 62 years old were categorized as "young". Missing data in "age_at_initial_pathologic_diagnosis" was excluded. Also, stage '– and stage X were excluded since they have no meaning.

The visualizations and analyses were created using an assortment of libraries within R. In R, TCGABioLinks was used to access the data. Survival and Survminer were used to create Kaplan-Meier Curves, ggplot2 was used to create bar plot, and plotly was used to create interactive plots.

Reference

Breast cancer statistics: How common is breast cancer? American Cancer Society. (n.d.). Retrieved December 5, 2022, from https://www.cancer.org/cancer/breast-cancer/about/how-common-is-breast-cancer.html

Stage 1. Stage 1 | Breast cancer | Cancer Research UK. (2020, July 10). Retrieved December 5, 2022, from https://www.cancerresearchuk.org/about-cancer/breast-cancer/stages-types-grades/number-stages/stage-1#:~:text=Stage%201%20breast%20cancer%20means,breast%20cancer%20is%20very%20complex.

Stage 2. Stage 2 | Breast cancer | Cancer Research UK. (2020, July 10). Retrieved December 5, 2022, from https://www.cancerresearchuk.org/about-cancer/breast-cancer/stages-types-grades/number-stages/stage-2#:~:text=Stage%202%20breast%20cancer%20means,decide%20which%20treatment%20you%20need.

Stage 3. Stage 3 | Breast cancer | Cancer Research UK. (2020, July 10). Retrieved December 5, 2022, from https://www.cancerresearchuk.org/about-cancer/breast-cancer/stages-types-grades/number-stages/stage-3#:~:text=Stage%203%20means%20that%20the,decide%20which%20treatment%20you%20need.

Treatment of stage IV (metastatic) breast cancer. American Cancer Society. (n.d.). Retrieved December 5, 2022, from https://www.cancer.org/cancer/breast-cancer/treatment/treatment-of-breast-cancer-by-stage/treatment-of-stage-iv-advanced-breast-cancer.html#:~:text=Stage%20IV%20cancers%20have%20spread,the%20brain%20or%