The National Breast Cancer Foundation is an Australian non-profit organisation dedicated in funding life-saving breast cancer research. NBCF has invested in innovative research to improve prevention, diagnosis, treatment of breast cancer patients, with the goal of achieving zero deaths from breast cancer by 2030.

Recommendation

To increase the predictability of cancer patients, NBCF should conduct analysis on the influence of demographic variables such as age and race on breast cancer outcomes. This can aid in identifying high-risk population and enable precise treatments while spreading awareness. Resulting in a likelihood in reducing breast cancer mortality.

Evidence

The data analysed in this research is sourced from SEER Breast Cancer Data (IEEEDataPort), contributed by JING TENG. The dataset involves 4024 female patients, including various features on their demographic and cancer status.

1. Analyse survival months by Age groups, Race, and Marital Status.

Survival Months by Age groups, Race, and Marital Status.

Comparing the impacts of these factors through boxplot shows noticeable patterns hinting at possible relation between unique features leading to fewer survival. Most remarkable is the separated marital status, having a lower median at 67 compared to 73 averaged from other marital status. This result suggests some demographic groups would undoubtedly impact breast cancer mortality rate.

1. Evaluate survival curves by Age Group

The survival curve displays 2 trends, a faster decline rate for age 31-40 and 60-70, and a slower decline rate for 41-60. Similar study conducted in 2016 shows matching results – “Middle-aged breast cancer patients showed better survival than younger and older groups, except in advanced-stage disease” [1]. These replicable results implying the susceptibility of low and high aged breast cancer patients, showcasing mortality follow an age-based trend, and that possibility of other demographic factors contributing to fatality.

1. Positive Regional Node Percentage vs Age Group

Highlighted by the density graph of Positive Regional Node Percentage (Regional Node Positive / Regional Node Examined) against age, we can found hot spots occurring denser as age increase. Most apparent at around age 47 and reduce slightly as age increase. Research conducted in the United States shows - Women aged 40-49 show higher breast cancer incidence rates and significant annual increases compared to other age groups [2]. Proving that patient’s near age 50 grow more cancer cells than any other age group, marking them more prone to be being diagnosed with breast cancer.

~~The regression line annotated on the graph follows the equation~~

~~Positive Cancer Cell Percentage = 0.2597 + 0.0012 x Age~~

~~The equation displays a slightly sloped line near 0.26, suggesting the increasing nature of the severity of cancer percentage as the sample age. This is due to a higher proportion of elderly patients having more percentage of cells tested to be positive, which follows the known tumour behaviour to grow. To reduce mortality rates in breast cancer patients, it would be critical to also necessary care and medication to the older generations.~~

Limitations:

The dataset recorded is cut-off data, any surviving patient's survival month is not accurately recorded and may impact the calculations and graphs.

It is possible that the cause of death may not be directly a result of breast cancer. This can cause false entries that ends abruptly.

Few entries of age 21-30. No defined trend can be observed, and low numbers of data can produce unrepeatable experiments. To prevent this, no conclusions were made with this age group specifically.

The sample is collected regionally, meaning that the data may not function in other countries due to difference in temperature, humidity and such.

Conclusion:

This report highlights the predictability of breast cancer patients by analysing graphs and trends. National Breast Cancer Foundation should divert more resources into understanding influence of demographics including but not limited to age and race, supporting reduction of breast cancer mortality by identifying high-risk groups.

Acknowledgement

* [1] Chen HL, Zhou MQ, Tian W, Meng KX, He HF. Effect of Age on Breast Cancer Patient Prognoses: A Population-Based Study Using the SEER 18 Database. PLoS One. 2016; 11:e0165409. 10.1371/journal.pone.0165409 [[DOI](https://doi.org/10.1371/journal.pone.0165409)] [[PMC free article](https://pmc.ncbi.nlm.nih.gov/articles/PMC5087840/)] [[PubMed](https://pubmed.ncbi.nlm.nih.gov/27798652/)] [[Google Scholar](https://scholar.google.com/scholar_lookup?journal=PLoS%20One&title=Effect%20of%20Age%20on%20Breast%20Cancer%20Patient%20Prognoses:%20A%20Population-Based%20Study%20Using%20the%20SEER%2018%20Database.&author=HL%20Chen&author=MQ%20Zhou&author=W%20Tian&author=KX%20Meng&author=HF%20He&volume=11&publication_year=2016&pages=e0165409&pmid=27798652&doi=10.1371/journal.pone.0165409&)]
* [2] Brinton LA, Sherman ME, Carreon JD, Anderson WF. Recent trends in breast cancer among younger women in the United States. J Natl Cancer Inst. 2008;100(22):1643–1648. doi: 10.1093/jnci/djn344. [DOI] [PMC free article] [PubMed] [Google Scholar]
* <https://academic.oup.com/jnci/article/100/22/1643/2519082>
* <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0165409>