A study of prognostic factors for Breast Cancer

T.D.R. Peiris and M.R. Sooriyarachchi

Department of Statistics, Faculty of Science

Introduction

Background

Cancer is a group of diseases. It occurs when cells become abnormal and divided without control or order. It affects persons all over the world. Breast cancer (malignant breast neoplasm) is cancer originating from breast tissue, most commonly from the inner lining of milk ducts or the lobules that supply the ducts with milk. Cancers originating from ducts are known as ductal carcinomas; those originating from lobules are known as lobular carcinomas.

Very rarely men also get breast cancers. Men possess a small amount of nonfunctioning breast tissue (breast tissue that cannot produce milk) that is concentrated in the area directly behind the nipple on the chest wall. Like breast cancer in women, cancer of the male breast is the uncontrolled growth of the abnormal cells of this breast tissue.

The size, stage, rate of growth, and other characteristics of the tumor determine the kinds of treatment for breast cancer patients. Treatment may include surgery, drugs (hormonal therapy and chemotherapy), radiation and or immunotherapy. Some breast cancers are sensitive to hormones such as estrogen and or progesterone, which make it possible to treat them by blocking the effects of these hormones.

Rationale

Breast cancer is the most commonly prevalent type of cancer among women in Sri Lanka. Therefore it is worthwhile studying about breast cancer.

Objective

The objective of this study was to investigate multiple aspects related to breast cancer patients in Sri Lanka. The main objective was to determine the Influential prognostic factors that affect the survival time of breast cancer patients.

Details of data

Data for this study was collected from breast cancer patients registered during the year 2006 to 2008 at General Cancer Hospital Maharagama. There were 2907 patients suffering from breast cancer, among which were 2833 female patients and 74 male patients during this period.

In the hospital record room they measured patient identity number, age, sex, district, province, type of cancer by ICD code, survival status, registered date and date of last visit or date of death.

Methodology

The Kaplan-Meier survival curves and hazard plots were used to descriptively study the effects of different factor levels of each independent variable on survival. The log-rank test was used to confirm associations at a univariate level. For the advanced analysis the Cox proportional hazards model was used for modeling (collet 1994).

Outcome and Conclusions

The log-rank test indicated that age category, gender, race, menopausal status, and ward or clinics are the most significant variables that affect the survival time. Kaplan-Meier plots show that pre-menopausal women have higher survival time than post-menopausal women. Males have lower survival time than females.

In this study significant difference cannot be seen among breast cancer patients according to districts and provinces. This may be due to the relationship between genetic affect and breast cancer. Sinhalese patients have a higher survival time than other races. Very young and middle aged patients have a higher survival time while teenagers and very old patients have very low survival time.

The hazard plots indicate that the hazard of death increased very sharply for about two and half years for very old people. During the first year very young people have a high risk. Females have a lesser hazard when compared to males. Hazard of death increased very sharply for about three years for males when comparing with females. Pre-menopausal females have a lesser hazard with respect to the post-menopausal females. For both the pre and post-menopausal patients hazard of death increased very sharply for about three years.

The Cox proportional hazard regression model is used to model failure time or follow up data. This model is based on the assumption of proportional hazards, and no particular form of probability distribution is assumed for the survival times. In this study for survival data modeling Cox-proportional hazard model indicates that ward or clinic, age category, sex, menopausal status and ethnicity affect the survival time. The selection of the variables for inclusion in the model based on the deviance step wise selection method was used.

The model obtained as follows

$$h_i(t) = \exp\{ \text{agecategory}_p + \text{ward or clinic}_q + \text{sex}_r + \text{menopausal \&nonmenopausal}_s + \text{race}_a \} h_0(t)$$

Table 1 gives the hazard ratios for different variables.

Table 1- Results of Parameter Estimates, P-values, Hazard Ratios

Variable	Parameter estimate	P-value	Hazard Ratio
agecategory			
agecategory1	-0.978	0.106	.376
agecategory2	-1.500	.000	.223
agecategory3	-.984	.000	.374
Sexcode	.038	.873	1.038
Mennon			
mennon0	.844	.000	.525
mennon1	-0.432	0.003	0.346
Racecode			
Racecode1	-.404	.186	.667
Racecode2	-.217	.505	.805
Racecode3	-.150	.639	.861

Interpretation of Coefficients and Hazard Ratios

A positive sign in regression coefficient means that the hazard is higher at any given time for the associated level of the factor compared to the baseline level of the factor. Negative sign in regression coefficient means that the hazard is lesser at any given time for the associated level of the factor compared to the baseline level of the factor.

Patients who are in ward have about 0.66 times lesser risk of death than patients who are in clinic at any given time. This is highly significant.

Very young, teenage and middle aged patients have lesser risk of death than very old patients. All these age groups are having negative parameter estimates compared to very old patients.

Postmenopausal women have about 0.525 times more risk of death than male patients while pre-menopausal women have about 0.346 times less risk of death than males at any given time.

Sinhalese, Tamils and Muslims have lesser risk compared to other race category.

Recommendations

The Government can conduct public health awareness programmes on breast cancer, because the early detection will pave the way to complete curability for breast cancer patients.

References

Collet D,(1994), “Modeling Survival Data in Medical Research”,Chapman& Hall, London.