Novel Approaches for visualising Time to Event Health Data

Victor Pacifique Rwandarwacu, Ranganath Shyamsundar[1]

[1] School of Mathematical and Statistical Sciences, University of Galway

Definition and Background

Time-to-event or survival Analysis is the analysis of data in the form of times from a well-defined time origin until the occurrence of some particular event or end point1[1]. Survival data are generally asymmetric and censored; which imply use of specific approaches for analysis and visualizations, such as survival function and Kaplan Meier(KM) plot.

Survival function $S(t)$ is the probability that the survival time is greater than or equal to time (t) which is the observed value of random Variable T with distribution function $F(t)$[2].

$$S(t) = \mathrm{P}(T \geqslant t) = 1 - F(t)$$

$$F(t) = \mathrm{P}(T < t) = \int_0^t f(u)\mathrm{d}u$$

The life table estimate of the survival function at Jth interval is given by:

$$S^*(t) = \prod_{i=1}^{j-1} \left(\frac{n_i' - d_i}{n_i'} \right)$$

for $t_{j-1}' \leqslant t < t_j', j = 2, 3, \ldots, m.$ d_j and c_j denote the number of deaths and the number of censored survival times, respectively, in this interval, n_j = the number of individuals at risk of death, at the start of the j th interval. n_i' =number of individuals at least in interval j

Survival Ratio plot is a robust approach for comparing survival distributions [3],

$$R(t) = \frac{S_1(t)}{S_2(t)}$$

This Project will explore novel informative visualization of time to events data specifically comparing survival curves of different covariates or treatment in the trial.

Objectives of Project

1. Explain the statistical standard concept and visualization in time to events data.
2. Propose novel Visualization of survival data in health sector.

Data Sources and Datasets

The dataset is from NIH National Cancer Institute , TCGA Program on a project called “Breast invasive carcinoma (BRCA)”, it contains information about: demography, exposure , Family History(regarding cancer), Follow up, Molecular Test, other Clinical Attribute, pathology detail,and Treatment of female Breast cancer patients diagnosed and followed up for different outcomes.For demonstration, our analysis focuses on Survival outcomes by pathologic stages [4]

Proposed Approach

Table 1 demonstrate the pattern of survival function and the change on the number of people at risk on each time interval exported from a survival model.

Table 1: BRCA data survival

Time	Survival	n.risk	Std.Error	Lower.95CI	Upper.95CI
5.256673	0.1868138	233	0.0182546	0.7781836	0.8497633
5.275838	0.1903494	230	0.0185144	0.7741642	0.8467637
5.456537	0.1939965	222	0.0187868	0.7700105	0.8436791
5.500342	0.1976601	220	0.0190553	0.7658481	0.8405705
5.741273	0.2014991	209	0.0193470	0.7614679	0.8373351
5.823409	0.2053942	205	0.0196408	0.7570282	0.8340488

Figure 1,Figure 2, Figure 3 and Figure 4 highlight different approach of visualizing data from standard KM plot, KM with covariates, survival ratio with confidence interval as well as permutation envelopes respectively.

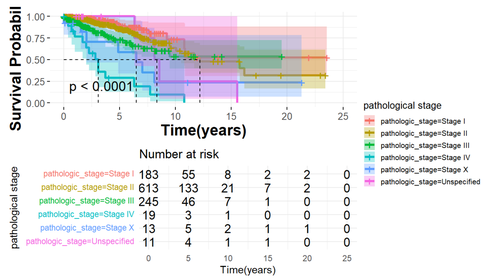

Figure 1: KM plot_all pathologic stages

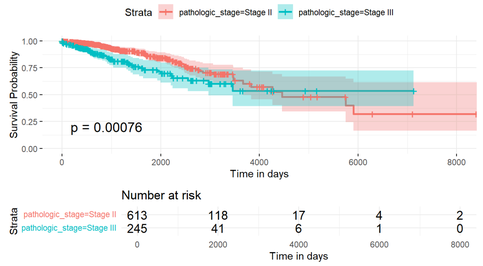

Figure 2: KM plot Of Pathologic stage II and III

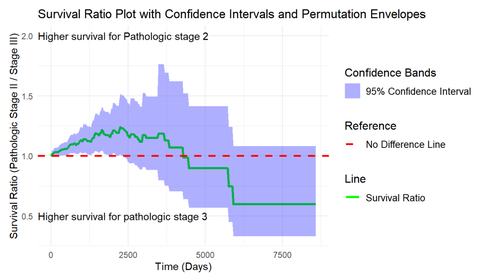

Figure 3: Survival Ratio plot for Path. stage II/ III with C.I

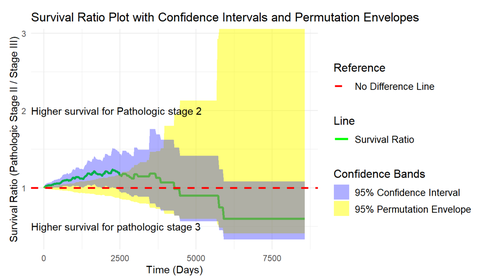

Figure 4: Survival Ratio plot for Path. stage II/ III with C.I and Permutation envelope

Next Project Steps

- Explore different visualization approaches for parametric non parametric survival Methods.
- Extensive literature review and feedback collection on the generated plots.

GitHub

The code and data sets for this project can be viewed at our GitHub repository here: https://github.com/rwandarwacu1/Msc_thesis_survival

References

1. David Collett, Modelling survival data in medical research , Fourth Ed.↩
2. Peace, Karl E.. Design and Analysis of Clinical Trials with Time-to-Event Endpoints (Chapman & Hall/CRC Biostatistics Series) (p. 74). CRC Press. Kindle Edition.↩
3. J. Newell et.al https://doi.org/10.1016/j.compbiomed.2005.03.005↩
4. <TCGA-BRCA , https://portal.gdc.cancer.gov/projects/TCGA-BRCA>↩