Novel Approaches for visualising Time to Event Health Data

Ranganath Shyamsundar[1]

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

OLLSCOIL NA GAILLIMHE
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 requires the use of specific approaches for analysis and visualisations, such as this survival function, Kaplan Meier(KM) estimator and plot.

The 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 Kaplan Meier estimate of the survival function at kth interval is given by:

$$\hat{S}(t) = \prod_{j=1}^{k} \left(\frac{n_j - d_j}{n_j} \right)$$

For $t_{(k)} \leqslant t < t_{(k+1)}, k = 1, 2, \ldots, r$, with $\hat{S}(t) = 1$ for $t < t_{(1)}$,where t_{r+1} is taken to be ∞ d_j denotes the number of deaths and the number of censored survival times, respectively, in this interval, n_j is the number of individuals alive just before $t_{(j)}$ and d_j deaths at $t_{(j)}$.

Survival Ratio, a robust approach for comparing survival distributions [3], is defined by:

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

This project explores the use of novel informative visualisations of time-to-event data, specifically comparing survival curves of different covariates or treatments in a trial.

Data Source and Dataset

The dataset originates from a clinical research study focused on dialysis patients, providing comprehensive information on various factors influencing patient outcomes. It includes details on demographics, medical history, comorbidities, treatment initiation, and survival outcomes. Key variables encompass age, dialysis initiation time, presence of comorbid conditions (such as diabetes, hypertension, and renal disease), and follow-up data on survival status. For demonstration purposes, this analysis centers on survival outcomes, comparing groups based on the presence or absence of diabetes.

Proposed Approach

Table 1 demonstrates the section of survival function and the change of the number of people at risk on each time interval.

Table 1: **Dialysis Data Survival**

Time_Yrs	Survival_Prob	n.risk	Std.Error	Lower.95CI	Upper.95CI
1	0.9663483	6805	0.0021860	0.9620732	0.9706423
2	0.9274938	6168	0.0032007	0.9212416	0.9337884
3	0.8998527	5503	0.0037634	0.8925067	0.9072591
4	0.8804762	5062	0.0041146	0.8724485	0.8885777
5	0.8640583	4773	0.0043905	0.8554958	0.8727066
6	0.8518617	4534	0.0045856	0.8429213	0.8608969
7	0.8414251	4326	0.0047482	0.8321700	0.8507831
8	0.8274115	4143	0.0049598	0.8177473	0.8371899
9	0.8157189	3892	0.0051342	0.8057178	0.8258440
10	0.8055775	3700	0.0052836	0.7952881	0.8160001

Figure 1, ?? and 2 highlight different approaches of visualising the estimated survival function.

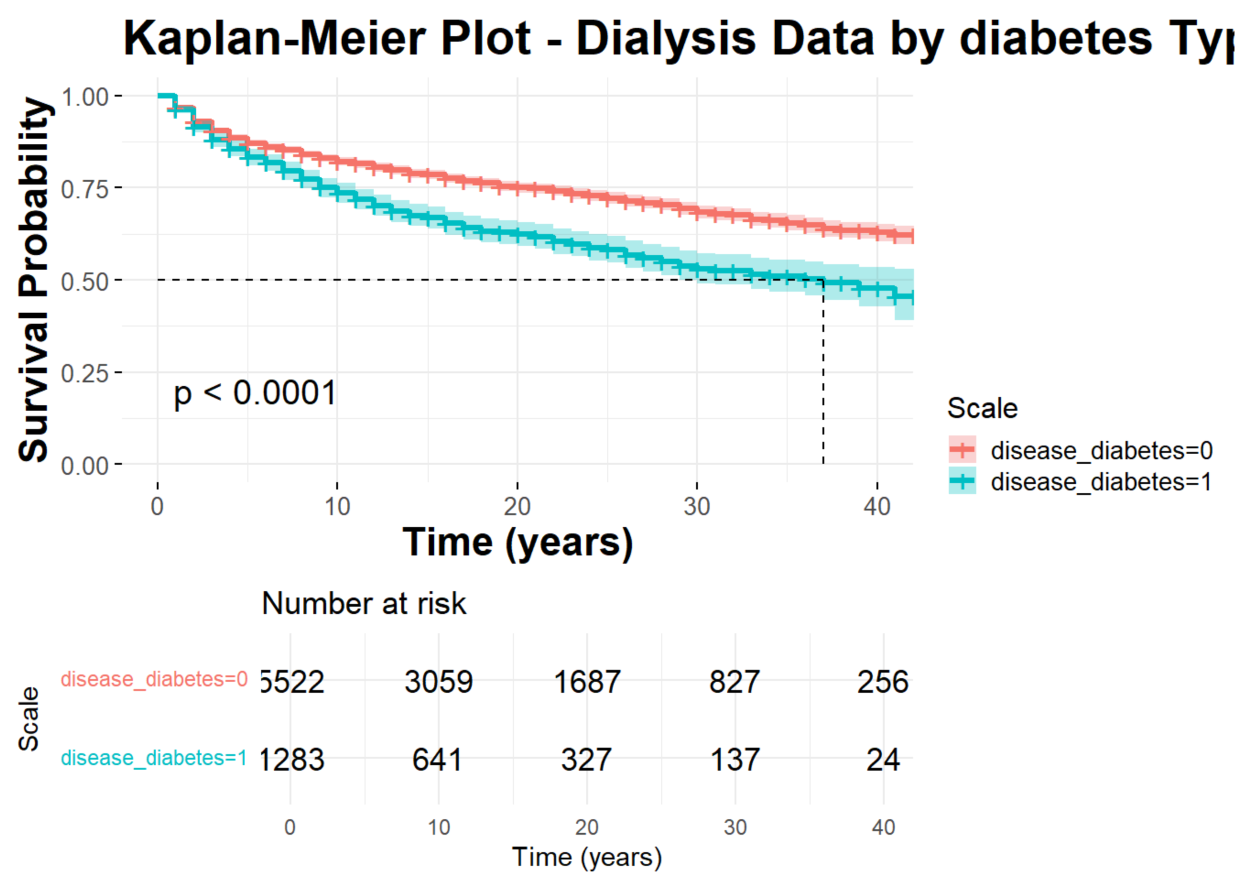

Figure 1: *KM plot for diabetes vs no diabetes*

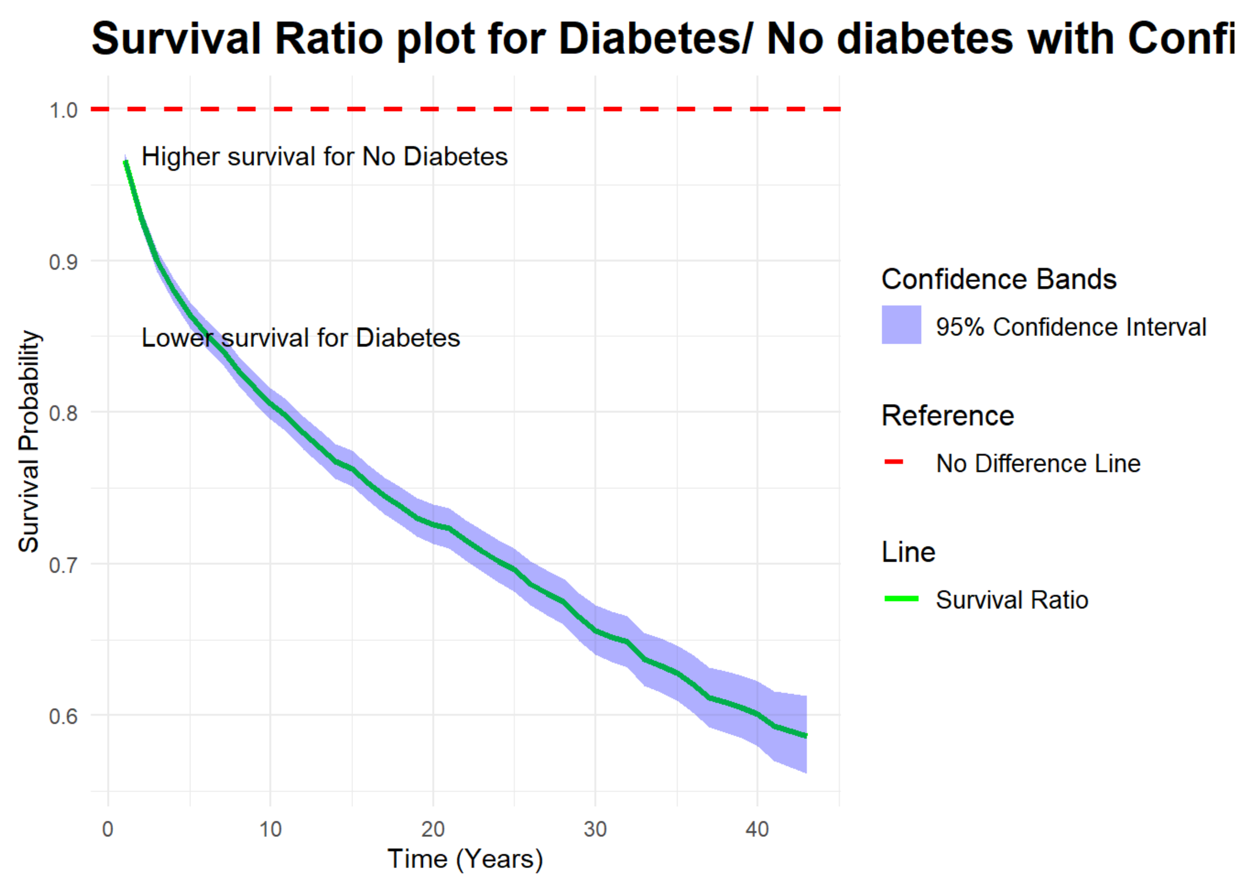

Figure 2: *Survival Ratio plot for diabetes vs no diabetesI with 95% C.I*

Next Steps

-Visualize survival differences between independent groups, incorporating confidence intervals to assess variability and significance.

-Generate survival ratio plots for paired data, using permutation envelopes as reference bands to evaluate deviations and provide robust comparisons.

-Compare survival distributions across more than two groups utilizing non-parametric statistical methods to identify significant differences.

GitHub

The code and dataset for this project can be at GitHub repository through this link : 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, Survival ratio plots with permutation envelopes in survival data problems, https://doi.org/10.1016/j.compbiomed.2005.03.005↩