Survival Analysis in the International Strokes Trial Database: Comparison of Heparin and Aspirin to Address Stroke Over a 6-month Period

By: Maruthi Kumar Mutnuri & Atlanta Liu

DATA 621: Advanced Statistics Modelling

University of Calgary

Winter 2020

Introduction

Background

Cerebrovascular diseases, or more commonly known as strokes, are the third leading cause of pre-mature deaths in Canada [1]. Strokes happen when there is a blockage (ischemic) or rupture (hemorrhagic) in the blood vessels that lead up to the brain, causing sudden impairment in one's ability to properly function. If blood supply is reduced in a brain region that controls vital organs or one's ability to breath, then it will generally lead to one's death if the individual is not given prompt treatment.

For patients suffering from ischemic stroke, there are several medications available to treat its symptoms. More commonly used medications tend to involve blood thinners, such as anti-platelet drugs or anticoagulants. Aspirin falls into the anti-platelet category, which helps to prevent platelets from clotting up together, thereby preventing excessive blood vessel blockage to the brain. On the other hand, Heparin is also widely used and belongs to the family of anti-coagulants. Heparin is similar in that it also functions to prevent blood clotting, however it does so by blocking production of certain proteins that lead up to the blood clotting.

Problem Definition & Hypothesis

While both types of drugs have been widely used, there is still some debate on the effectiveness of certain drugs like Heparin when compared to Aspirin to address re-occurring ischemic strokes. At times, physicians find themselves in a situation where they could prescribe either drug (or even both drugs at once) to the patient, but not know which type of drug would be most effective for the patient. To address this issue, the International Stroke Trial (IST-3) has been designed to test the effectiveness of both drugs. This report will analyze the results of their randomized clinical trial through a survival analysis. By looking at the survival rates between patients who were allocated either Aspirin or Heparin, we can determine which drug improves overall survivability across a 6-month period the most.

Patients who only take Aspirin are expected to show an improvement in survival rates compared to our baseline group (no aspirin and no heparin) [3]. For patients who take a high heparin dosage only, we expect the survival outcomes to show similarities to the baseline group as described in a related research [3].

Guiding Question

Is there a significant difference in survival rates between individuals who were randomized to take either Aspirin or a Heparin dosage?

- For individuals who took both heparin and aspirin, do we see an even greater increase in survival rates when compared to taking only one type of medication?
- How does age and sex impact the survival rates between patients who take Aspirin only or Heparin only?

Methods

Data Collection

The dataset used in this report was collected from the IST database in a comma-separated values (CSV) format, available online for public use at the University of Edinburgh. The IST dataset consisted of 19,345 patients with symptoms of stroke, who were then randomized into multiple groups to test the safety and efficacy of aspirin and heparin [2]. These patients were selected based on the physician's criterion. If the physicians were uncertain to administer either aspirin or heparin (or both) to the patient, they were eligible for enrollment to the trial [2]. In terms of anonymization, the data custodians have removed all direct and indirect identifiers to the patients [4]. Dates and events have been converted to the number of days from randomization to help further reduce the chances of identifying the patients.

Variable Definitions

Patients were considered in the "Aspirin only" group if they had only taken Aspirin (RXASP = Y) and no Heparin (RXHEP = N). Likewise, individuals were considered in the "High Heparin" group only if they received a medium or high dosage of heparin (RXHEP = M or H) and no Aspirin (RXASP = N). Note that medium is included together with high dosages because the original pilot study coded them to be the same amount despite wording it differently. Individuals who received both Aspirin and a high dosage of Heparin are included separately to determine if there is an additive effect when combining both medications. Finally, patients who did not receive either Aspirin or Heparin (RXASP = N, RXHEP = N) are used to measure the baseline effect.

In the IST dataset, Heparin is provided to the patients as a subcutaneous injection. The dosages amount provided to the patients could be either low (5000 IU) or medium/high (12,500 IU). On the other hand, individuals who were allocated to receive Aspirin would be instructed to intake 300mg daily. The event of interest is death, which will be analyzed across a six-month period from the start of each patient's randomization date to determine the overall survival probability of each medication group. Patient compliancy and research attrition did not pose any issues during the trial [2].

The two main variables of interest for our survival analysis is time of death (TD) as well as indicator for death (DIED: 0 = did not die, 1 = died). Time of death is measured in days, and censoring is accounted by the status of the patient. Several patient characteristics are used as explanatory variables in the modelling process. This includes age in years, sex, atrial fibrillation (RATRIAL), systolic blood pressure (RSBP) in mmHg, as well as any visible infarcts on the patients CT scan (RCT). Note that there were 984 patients who did not have their atrial fibrillation status recorded during the pilot study, so these rows were removed from our data frame. There are multiple types of subgroups for stroke, with 17398 (90%) being diagnosed with ischemic stroke, 599 hemorrhagic stroke, 992 unknown, and 420 without stroke in the

dataset. This analysis will include all subtypes for strokes. These characteristics were specifically used in our analysis as they were measured before randomizing each patient to their respective treatment group.

Procedures

The survival analysis consists of a Kaplan-Meier plot, log-rank test, as well as a Cox Proportional Hazards (CoxPH) model, which were all analyzed in R. The Kaplan-Meier survival curve is used to visualize the survival probability of each treatment group across the six-month period. Confidence intervals were not included in the plot as the survival curves between all four groups were quite similar. Since we are only interested in the first 6 months, we limited our analysis to only include the first 180 days. Any patients who survived past 180 day were marked as censored. The alpha level used for all testing procedures is 0.05.

The proportionality assumption between each group were then determined by pairing the survival curve with the Schoenfeld residuals plot and visually inspecting it for overlapping curves and checking the covariates were non-proportional ($p > 0.05$). Martingale plots are also provided for continuous covariates to test if there are systematic departures from linearity over time.

Following this, we conducted a log-rank test for all the treatments in comparison to our control group (no heparin & no aspirin) to determine if there were any significant differences in survival rates. Finally, we placed each patient characteristic into the CoxPH model to determine which variables were the most significant predictors of survivability in patients suffering from stroke. The modelling process involved looking at potential interactions as well as accounting for confounding variables (Δ beta > 10%). Model selection occurred through multiple likelihood ratio tests (LRT) to find the most parsimonious model that accounted for the most variance in survival rates.

Results

Data Exploration

To visualize the distribution of the demographics in our dataset, we plotted a bar chart with the number of counts in each age group and gender (Figure 1). Across the 19435 patients, we see that the sample is largely representative of the older population between ages 70-79 and 80+ (n = 12177, 63%). Additionally, there is an overall even distribution of males (n = 10407, 54%) to females (n = 9028, 46%) in the dataset. The patients were then grouped into four treatments, control (n = 4860, 25%), aspirin only (n = 4858, 25%), low heparin only (n = 2429, 12%), high heparin only (n = 2426, 12%).

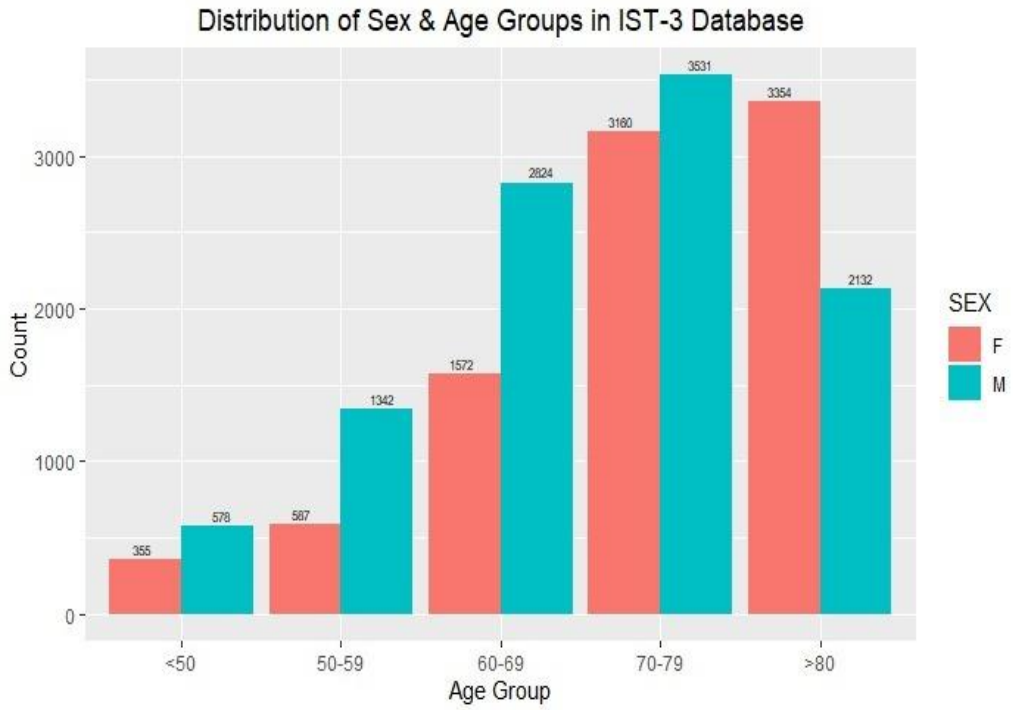

Figure 1: Distribution of age and sex in patients suffering from re-occuring stroke from IST-3 data

Kaplan Meier Estimator - Survival Curves

Figure 2 shows the Kaplan Meier Curve for all four treatment groups across time in days. A hazard table is provided below, showcasing the count of patients remaining as time increases. Looking at the graph, we see that all four groups appear to be very similar over the first 50 days, with a large overlap in curves. While the curves slowly separate apart from there onwards, there is still some overlap present when we compare low heparin with high heparin patients. This is an indicator that our model may not fully satisfy the assumption of proportionality even with our pre-specified time range. This assumption is further tested through a Schoenfeld Plot of Residuals, which is presented at a later stage in our modelling process.

Log-Rank Test

While a Log-Rank test may not be feasible due to the presence of overlap in our survival curves, we presented the analysis in Table 1 for reference. The null hypothesis (H_0) is that survival distributions between each group are equal at all follow-up times, whereas our alternative hypothesis (H_A) is that the survival curves differ at one or more points in time. The results of the log-rank test show that there is not a significant difference in survival between any two or more drug types ($\chi^2 = 5.8$, df = 3, p = 0.1).

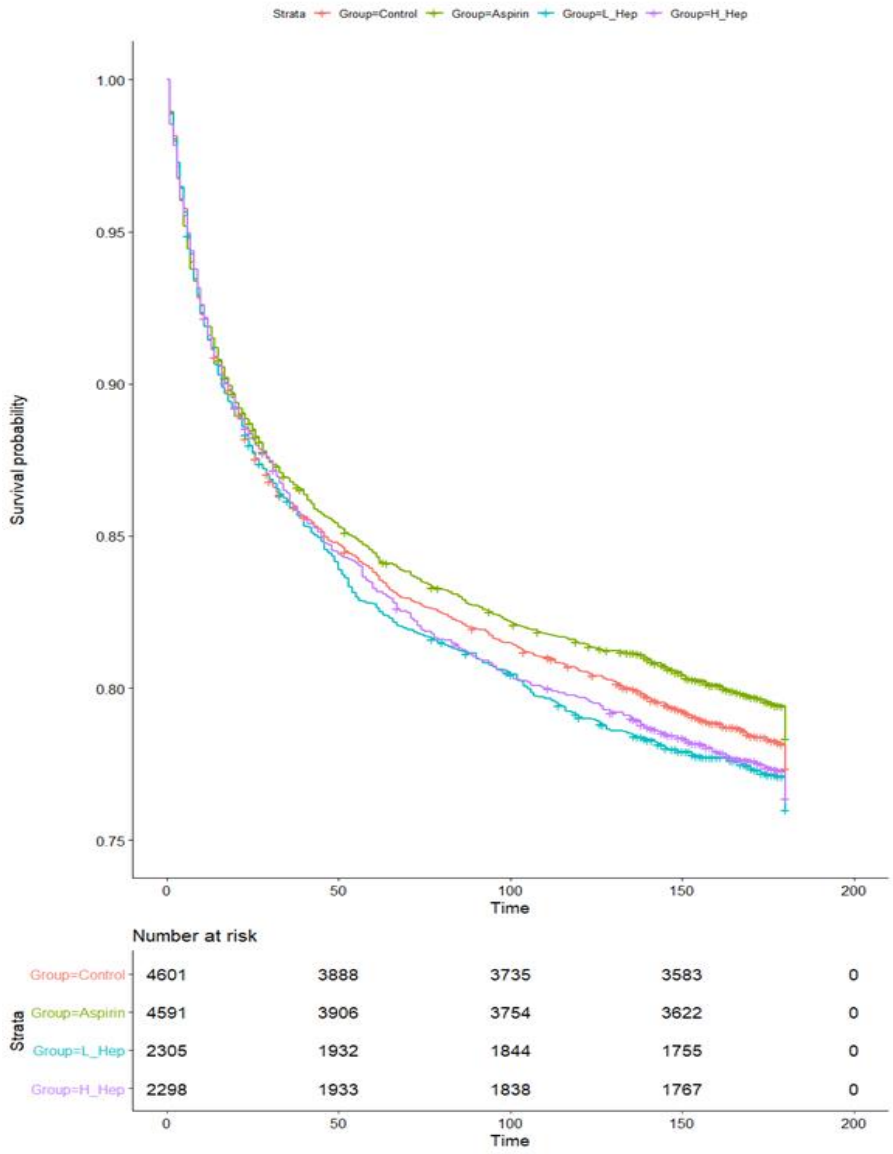

Figure 2: Kaplan Meier survival curve shows survival rates for each treatment across 180 days

```
```{r}
survdiff(Surv(TD, DIED)~Group, data = data)
```

Call:
survdiff(formula = Surv(TD, DIED) ~ Group, data = data)

                 N Observed Expected (O-E)^2/E (O-E)^2/V
Group=Control 4601   1030   1031   0.00108   0.00162
Group=Aspirin 4591   979   1033   2.77983   4.19296
Group=L_Hep   2305   545   513   1.97286   2.37600
Group=H_Hep   2298   536   513   1.01388   1.22105

 Chisq= 5.8 on 3 degrees of freedom, p= 0.1
```

Table 1: Log-rank test output comparing the survival rates for all four treatment groups. Results show that there is not a significant difference between any group.

Cox-Proportional Hazards Modelling

We began building our first cox-proportional hazards model by including only the different treatment groups as our base model. This model is used for subsequent comparison as we began our model selection process. The results of our full model are shown in the attached appendix. Looking at our base model, we see that the groups for each drug type are not significant predictors of survival on their own. More covariates will need to be added to see if there is an interaction with each drug type.

Base Model Equation: $log(h_i) = \beta_1 X_{Group}$

Variables	coef	exp(coef)	se(coef)	z	Pr(>|z|)
Group (Aspirin)	-0.04164	0.95922	0.04428	-0.94	0.347
Group (Low Heparin)	0.05416	1.05565	0.05281	1.026	0.305
Group (High Heparin)	0.04497	1.046	0.05297	0.849	0.396

Table 2: Summary of base model statistics

Our modelling process involves starting with the most complex model with all possible interactions and trimming down non-significant variables. We first removed any non-significant interactions based on the Wald Z test ($p > 0.05$), and subsequently test the reduced model through a likelihood ratio test. The new predictive variables added into our next model includes: Age, sex, systolic blood pressure, atrial fibrillation, and visible infarcts on the CT-scan.

Full Model:

$$log(h_i) = \beta_1 X_{Group} + \beta_2 X_{Age} + \beta_3 X_{Sex=M} + \beta_4 X_{RSBP} + \beta_5 X_{RATRIAL=Y} + \beta_6 X_{RVISFINF=Y} + \beta_7 X_{Group\times Age} + \beta_8 X_{Group\times Sex} + \beta_9 X_{Group\times RSBP} + \beta_{10} X_{Group\times RATRIAL=Y} + \beta_{11} X_{Group\times RVISINF=Y}$$

The interactions between Group:Sex, Group:RAtrial, and Group:RVISINF are removed since they showed no significant effect in the Wald Z test. This leaves us with the reduced model below. The two models are then put through a likelihood ratio test to see if we should adopt the more parsimonious model.

Reduced Model (1):

$$log(h_i) = \beta_1 X_{Group} + \beta_2 X_{Age} + \beta_3 X_{Sex} + \beta_4 X_{RSBP} + \beta_5 X_{RATRIAL=Y} + \beta_6 X_{\mathrm{RVISINF=Y}} + \beta_7 X_{Group\times Age}$$

Model	Log Likelihood	Chisq	Df	P(>|Chi|)
Full Model	-28739			
Reduced Model (1)	-28744	11.27	12	0.506

Table 2a: Likelihood ratio test results between full model and reduced model 1

Looking at the results of our first likelihood ratio test, we see that there is not a significant difference between the full model and the reduced model (χ^2 = 11.27, 0.506). This indicates that we should adopt the reduced model as it is more parsimonious. Next we try dropping the interaction between age and group, as the p-value for the Wald-Z test appears to be borderline significant (z = -1.823, p = 0.0682) in our first reduced model. We also looked to remove sex from our main effects as it was a non-significant predictor (z = 0.286, p = 0.7749). This leaves us with only the main effects in our new model. Again, we try the LRT to see if we should adopt the more parsimonious model. In this case it would be the second reduced model.

Reduced Model (2):

$$log(h_i) = \beta_1 X_{Group} + \beta_2 X_{Age} + \beta_3 X_{Sex} + \beta_4 X_{RSBP} + \beta_5 X_{RATRIAL=Y} + \beta_6 X_{RVISINF=Y}$$

Model	Log Likelihood	Chisq	Df	P(>|Chi|)
Reduced Model (1)	-28744			
Reduced Model (2)	-28747	5.3222	4	0.2558

Table 2b: Likelihood ratio test results between the first and second reduced model (removing Group*Age interaction)

The results of our second likelihood ratio test show that there is not a significant difference between the two models (χ^2 = 5.3222, p = 0.2558) when reducing the Group*Age interaction. This indicates that we can faithfully drop the remaining interaction term, leaving us with only the main effects. All remaining variables except for drug group show a significant effect through the Wald Z test ($p < 0.05$), indicating that there is no need to reduce the model further. While drug group could be reduced, this is part of our research question so it will remain in the model.

Final Cox-PH Model:

$$log(h_i) = \beta_1 X_{Group} + \beta_2 X_{Age} + \beta_3 X_{RSBP} + \beta_4 X_{RATRIAL=Y} + \beta_5 X_{RVISINF=Y}$$

Cox-Proportional Hazard Assumption Tests

Schoenfeld Plot of Residuals

The Schoenfeld Plot of Residuals is our main test to see if our final Cox-PH model passes the assumption of proportionality. Five plots are displayed for each of our main effects in the final model. We see that 4 out of the 5 variables meet the assumption of proportionality ($p > 0.05$). The only variable that does not is age ($p < 0.001$). This alone has a huge impact on our global Schoenfeld test, which is also revealed to be significant ($p < 0.001$). These results show that our model overall fails to meet the proportionality assumption. Hazard ratios and p-values can not be reliably trusted with this model, so they are not stated here.

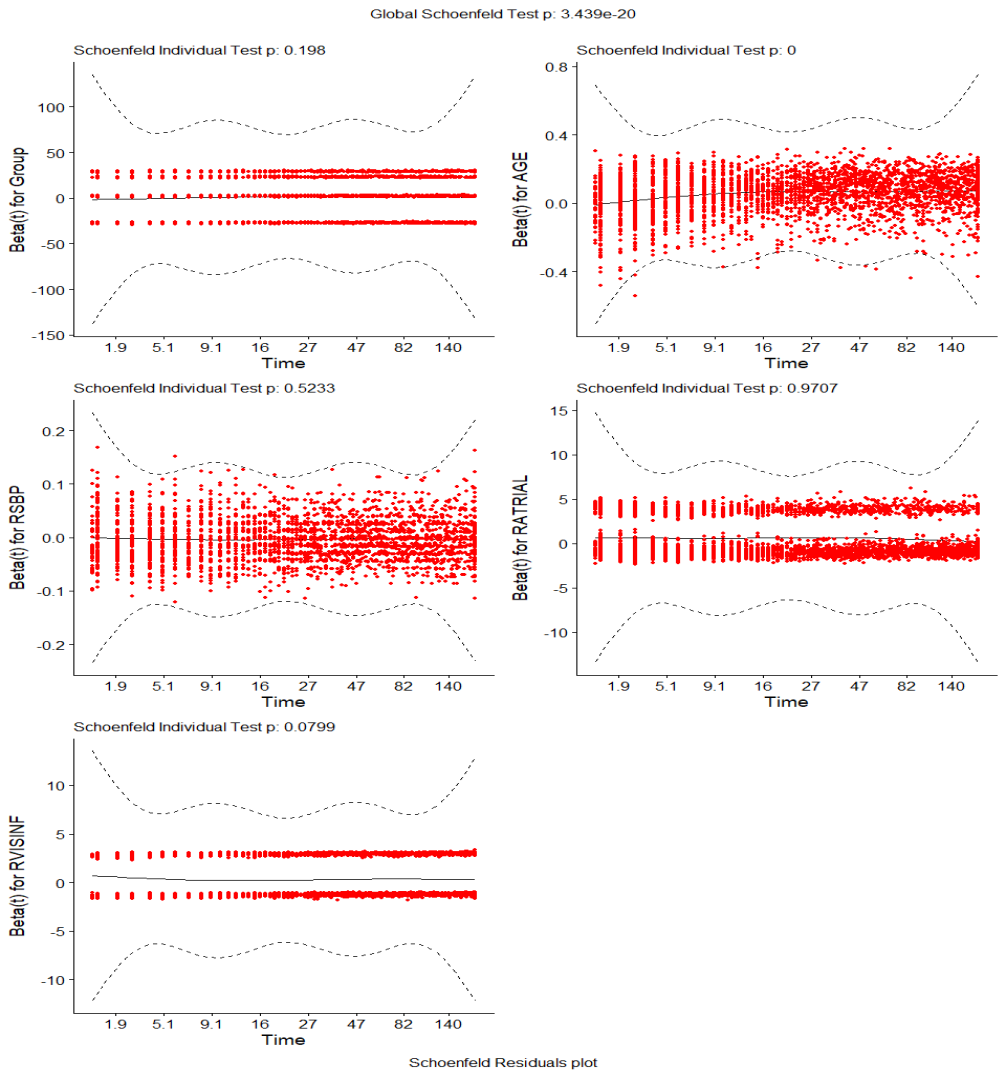

Figure 3: Schoenfeld residuals plot for the final Cox-PH model, results show that our model fails to meet the proportionality assumption due to accounting for a patient's age.

If we closely inspect the residuals on age, we see that the data points are quite sparse near the beginning, but they begin to become closely packed together as time increases. This indicates that the effect of age changes with time. Figure 4 shows this in further detail. There is a noticeable change in slope after 16 days for the effect of age. This gives us an indicator that age could be stratified across separate time intervals to meet the proportionality assumption.

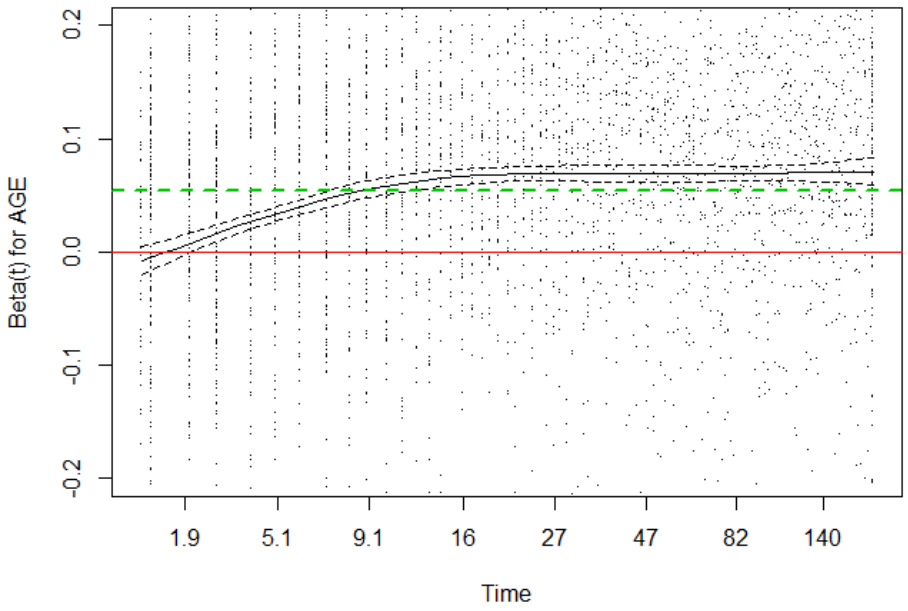

Figure 4: Beta residuals for age are displayed across time. Notice the change in slope from the beginning to approx. 16th day. This indicates that we can stratify these two time periods to meet the proportionality assumption.

Martingale Plot of Residuals

Figure 5 shows a martingale plot with multiple transformations on Age to see if we could reduce the non-linearity effects across time. This ultimately proved to be unsuccessful as there is still a large change in slope. A different option must be adopted to conduct a survival analysis.

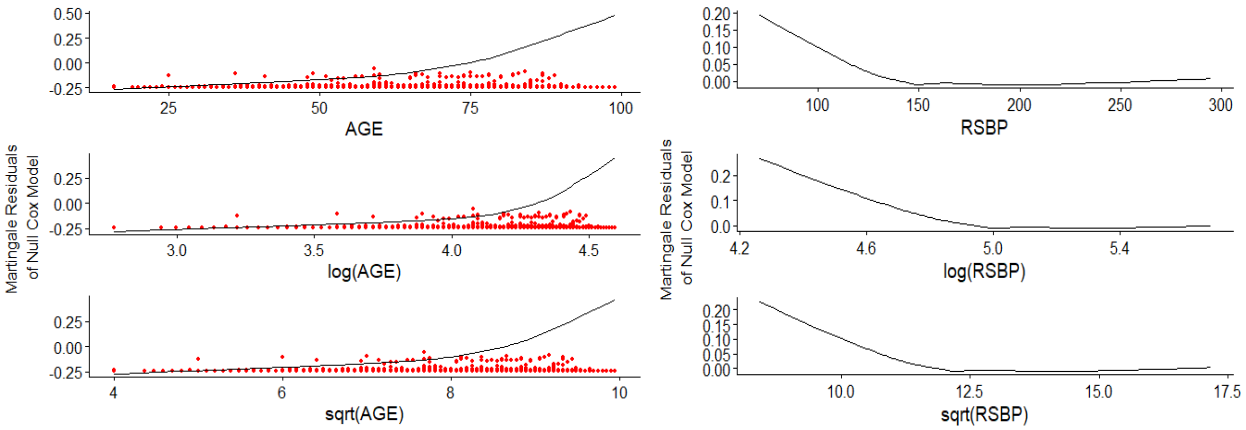

Figure 5: Martingale plot for our continuous variables. We applied a log and square-root transformations on age and systolic blood pressure. Still a visible change in slope, which suggests non-linearity in our covariates.

Splitting Survival Data at Specified Time Intevals

Using the survsplit() command, we split the timeline into 3 separate partitions by making a cut on day 4 and day 16. The time was then separated into days 0 to 4, 4 to 16, and 16 to 100. While we were initially interested in only the 16th day as our cut point for time, we were still unable to pass the proportionality assumption. As a result, we decided to add another time interval from 0-4 days, which helped us pass the test. Our final survsplit model uses the same variables as our most reduced Cox-PH model, except age is stratified by each time interval. We then have 3 separate beta/hazard ratios for age depending on the time interval of interest. The exact hazard probabilities are shown in our appended excel file for each group.

Final SurvSplit Model:

$$log(h_i) = \beta_1 X_{Control} + \beta_2 X_{RSBP} + \beta_3 X_{RATRIAL=Y} + \beta_4 X_{RVISINF=Y} + \beta_5 X_{AGE:tgroup=1} + \beta_6 X_{AGE:tgroup=2} + \beta_7 X_{AGE:tgroup=3}$$

Variables	coef	exp(coef)	se(coef)	z	Pr(>|z|)
GroupAspirin	-0.073757	0.928898	0.044645	-1.652	0.09852
GroupL_Hep	0.05114	1.052471	0.052977	0.965	0.33438
GroupH_Hep	0.058118	1.05984	0.053268	1.091	0.27525
RSBP	-0.002975	0.99703	0.000669	-4.446	8.74E-06
RATRIALY	0.550428	1.733996	0.040332	13.647	2.00E-16
RVISINFY	0.339934	1.404854	0.036961	9.197	2.00E-16
AGE:strata(tgroup)tgroup=1	0.013189	1.013276	0.00406	3.249	0.00116
AGE:strata(tgroup)tgroup=2	0.055798	1.057384	0.00387	14.419	2.00E-16
AGE:strata(tgroup)tgroup=3	0.069314	1.071773	0.002766	25.062	2.00E-16

Table 3a: Survsplit model summary statistic with three separate time intervals to stratify age.

Schoenfeld Plot of Residuals for Survival Split

Looking at our final plot Schoenfeld plot of residuals for the survsplit model, we see that our model barely passes the proportionality assumption (Global Schoenfeld Test: p = 0.05938). RATRIAL is the only variable that does not meet this condition, but it is kept in the model since it accounts for a significant amount of variance in survival rates. We can now interpret the hazard ratio for each coefficient much more reliably.

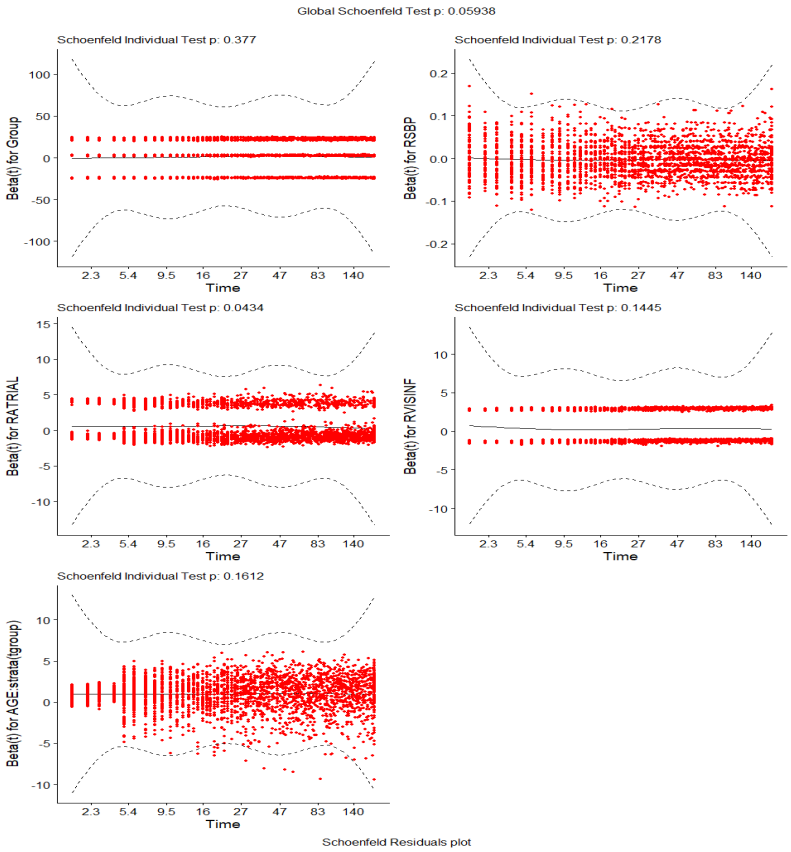

Figure 6: Schoenfeld residuals plot for Cox model with age stratified by time. Overall proportionality assumption is passed (p > 0.05) with this change.

Influential Outliers

Two different plots of residuals are shown in Figure 7, which presents the goodness of fit for our stratified Cox model. Both the covariates and the overall model show a fairly symmetric distribution around 0, indicating that there are no influential observations that need to be directly accounted for in our model.

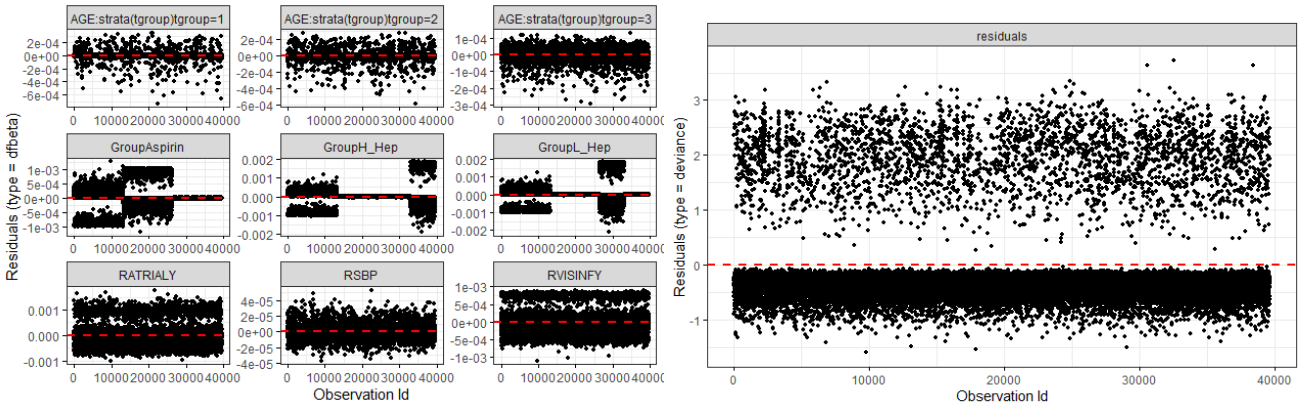

Figure 7: Plot of Residuals to check for influential outliers in our SurvSplit model

Hazard Ratio & 95% Confidence Intervals

Interpreting Coefficients

The hazard ratio for patients in the Aspirin group is 0.9289 ($CI_{95\%}$ = 0.8511 – 1.0138). This effect is borderline significant on its own (z = -1.652, p = 0.0985), further testing may be needed to validate its effects in a survival analysis. Since the hazard ratio is lower than 1, it indicates that patients in this group have a 0.9289 times lower risk of death than patients in the control group. But since it captures 1 within the confidence interval, there may not be a significant difference between these two groups.

The hazard ratio for patients in the low heparin group is 1.0525 ($CI_{95\%}$ = 0.9487 – 1.1676). Since the CI captures one, this indicates that there is not a significant difference in survival for patients in the low heparin group when compared to the control group.

The HR for patients in the high heparin group is 1.0598 ($CI_{95\%}$ = 0.9548 – 1.1765) The interval also captures one so there is not a significant difference in survival between high heparin and control treatments.

The HR for systolic blood pressure (RSBP) is 0.997 ($CI_{95\%}$ = 0.9957 – 0.9983). Since the hazard ratio is lower than 1, this indicates that as blood pressure increases there is a slight but significant decrease in risk of death. As blood pressure increases by 1 mm Hg then the odds of dying decreases by 0.3% (1-0.997 = 0.003 = 0.3%).

The HR for RATRIAL is 1.734 ($CI_{95\%}$ = 1.6022 – 1.8766). Since the HR is greater than 1, this indicates that the relative risk of death for patients with atrial fibrillation is 1.734 times higher than patients without it.

The HR for RVISINF is 1.4049 ($CI_{95\%}$ = 1.3067 – 1.5104). Since HR is greater than 1, this indicates that patients who have visible infarcts on their CT scan are at a much greater risk of death than individuals who do not have a visible infarct. Specifically, the relative risk is 1.4049 times more for those who do have a visible infarct in comparison to those who don't

Since the age variable did not meet the proportional hazards assumption, we stratified the age variable into 3 groups of time interval, so the interpretation of age will have following three coefficients.

- THE HR for age in the first time group (0 – 4 days) is 1.0133 ($CI_{95\%}$ = 1.0052 – 1.0214). This indicates that as age increases per year, their risk of death increases by about 1.0133 times more. As age increases by 1 year the odds of dying increases by 1.33% (1-1.0133 = 0.0133 = 1.33%).
- The HR for age in the second time group (4 – 16 days) is 1.0574 ($CI_{95\%}$ = 1.0494 – 1.0654). Again, as the age of a patient increases per year, their risk of death also increases in this time interval, albeit at a slightly higher rate than the first. Specifically,

the risk of death increases by about 1.0574 times more per increment in age. As age increases by 1 year the odds of dying increases by 5.74% (1-1.0574 = 0.0574 = 5.74%).

- The HR for age in the third time group (16 – 180 days) is 1.0718 ($CI_{95\%}$ = 1.066 – 1.0776). As the age of a patient increases per year, their risk of death is also increased in this time interval. The risk here is much greater than the first two time intervals, as the risk of death increases by about 1.0718 times more per increment in age. As age increases by 1 year the odds of dying increases by 7.18% (1-1.0718 = 0.0718 = 7.18%).

Variables	exp(coef)	exp(-coef)	lower .95	upper .95
GroupAspirin	0.9289	1.0765	0.8511	1.0138
GroupL_Hep	1.0525	0.9501	0.9487	1.1676
GroupH_Hep	1.0598	0.9435	0.9548	1.1765
RSBP	0.997	1.003	0.9957	0.9983
RATRIALY	1.734	0.5767	1.6022	1.8766
RVISINFY	1.4049	0.7118	1.3067	1.5104
AGE:strata(tgroup)tgroup=1	1.0133	0.9869	1.0052	1.0214
AGE:strata(tgroup)tgroup=2	1.0574	0.9457	1.0494	1.0654
AGE:strata(tgroup)tgroup=3	1.0718	0.933	1.066	1.0776

Table 3b: Hazard ratios and 95% confidence interval for each covariate in the Survsplit model.

Hazard function for each group (i = control, aspirin, low heparin, or high heparin):

$$log(h_{Group}) = \beta_1 X_{Group=i} + \beta_2 X_{RSBP} + \beta_3 X_{RATRIAL=Y} + \beta_4 X_{RVISINF=Y} + \beta_5 X_{AGE:tgroup=1} + \beta_6 X_{AGE:tgroup=2} + \beta_7 X_{AGE:tgroup=3}$$

Discussion

This survival analysis looked at examining the differences in survival rates between each type of drug treatment for patients suffering from stroke. Our initial Kaplan Meier curve showed that there was a large overlap between each group, despite pre-specifying our timeline to only include the first 180 days. This overlap between the survival curves provided a strong indicator that we could not properly run a log-rank test or a Cox-Proportional Hazards model due to the model not meeting the proportionality assumption, which was later confirmed through our Schoenfeld residuals plot. While this could have been presented as our final result, we decided to probe further into our model by stratifying age and splitting time intervals so that we could model which covariate had a significant contribution to the survival rates across the 6-month period.

This report did not specifically look at higher order since those covariates could not be properly presented in the Schoenfeld residuals plot, which was the crux of our proportionality

assumption testing. However, we did find that blood pressure had a significant higher order (cubic) term when accounted for. This model is shown in the portion of our attached R-appendix.

Our final time-split cox model revealed the most significant contributors to the risk of death in patients with stroke is systolic blood pressure, atrial fibrillation, visible infarcts, as well as age. We predicted that Aspirin would show a significant difference in survival rates when compared to the control group. However, our results only revealed a borderline significance. This might be reflective of how we cut the time intervals; a better method might be to stratify age using bins instead. When we look at the differences between low and high heparin in comparison to our control, it aligns neatly with our hypothesis that there is not a significant difference in the usage of heparin. Physicians may want to reconsider providing this medication to patients suffering from stroke patients in the future.

Study Limitations

Even though we have managed to address the proportionality assumptions in our cox model by splitting the time data, the results need to be carefully considered. Time intervals should be pre-specified in the hypothesis stage, rather than a means to address the proportionality assumption. The time intervals chosen were also based on visual inspection so that there was a greater chance of passing the assumption. Overall, there should have been a much more rigorous approach to address this issue, but we were unfamiliar with the different types of modelling techniques that could be utilized in this field.

We chose to include systolic blood pressure as a continuous variable rather than binning it as a categorical. This choice was predominantly made since categorization would have imposed discontinuities in each interval and assumption that the effect is the same across the entire interval. This translates into a loss of information. However, we later found that even though a rise in blood pressure can reduce risk of death, there is a U-shaped pattern where extreme ends of each BMI measures, low (<140) and high (>200), would lead to an increase in mortality. Binning these groups may have provided more accurate results and addressed the non-linearity issues we had.

Conclusion

Overall, our survival analysis revealed that there is not a large difference in survival rates in the first six months when comparing multiple treatment types for strokes. The presence of atrial fibrillation and visible infarcts negatively affect the survival rates no matter what group the patient is in. Further research in this area should look at survival analysis that does not rely on the assumption of proportionality, particularly in regard to variables such as age and RSBP.

References

1. Public Health Agency of Canada. Government of Canada [Internet]. , 2017 - Canada.ca. Government of Canada; 2019. Available from: https://www.canada.ca/en/public-health/services/publications/diseases-conditions/stroke-canada-fact-sheet.html

2. International Stroke Trial Collaborative Group. "The International Stroke Trial (IST): a randomized trial of aspirin, subcutaneous heparin, both, or neither among 19435 patients with acute ischaemic stroke," Lancet, vol. 349, pp. 1569-1581, 1997.

3. Hankey GJ, Berge E, Sandercock P. Anticoagulants Versus Antiplatelet Agents for Acute Ischemic Stroke. Stroke. 2003;34(6):1571–2.

4. Sandercock PA, Niewada M, Członkowska A. The International Stroke Trial database. Trials. 2011;12(1). Available from: https://trialsjournal.biomedcentral.com/articles/10.1186/1745-6215-12-101#Tab2

5. Schober P, Vetter TR. Survival Analysis and Interpretation of Time-to-Event Data. Anesthesia & Analgesia. 2018;127(3):792–8. Available from: https://journals.lww.com/anesthesia-analgesia/Fulltext/2018/09000/Survival_Analysis_and_Interpretation_of.32.aspx

6. Cox Model Assumptions [Internet]. STHDA. Available from: http://www.sthda.com/english/wiki/cox-model-assumptions