BS851 Spring 2025 Midterm Project
Reghan Lanning
11 April 2025

The International Stroke Trial (IST), conducted between 1991 and 1996 was a large, randomized controlled trial of individuals with suspected acute ischaemic stroke. The aim of the trial was to establish whether early administration of aspirin, heparin, both or neither influenced the clinical course of acute ischaemic stroke. For these analyses, we will use a subset of the data for participants randomized to Low Dose Heparin, High Dose Heparin or Heparin placebo in a 1:1:2 allocation. All participants in these groups also received aspirin.

In the same Blackboard folder as this midterm exam is a SAS dataset called IST_2025. The data consists of the following variables:

TRT: Randomized Treatment (a character variable, entered as "Low Dose", "High Dose", and "Placebo")
SEX: M=male, F=female
AGE: Age in years
RSBP: Systolic blood pressure at the time of randomization (mmHg)
STYPE: Stroke subtype (TACS, PACS, POCS, LACS, OTH=OTHER)
STROKE14: Recurrent stroke within 14 days. Y = yes, N=no (***Primary Outcome***)
ID14: Death within 14 days. Y=yes, N=no (***Secondary Outcome – time to event***)
TD: Time to death or censoring in days
CMPLASP: Compliance to aspirin Y=yes, N=no
CMPLHEP: Compliance to heparin Y=yes, N=no

PART 1: RANDOMIZATION

1. Imagine that you needed to create a randomization schedule for this study. The study was planned for 9,600 total participants from **10 countries** allocated 1:1:2 to dose 1, dose 2, and placebo. Use permuted block randomization using a block size of your choice and stratify by country. In addition, include your SAS code for this question in the body of your report and print the first 20 random allocations in one country. (Note that you do not need to look at the data to answer this question.) (5 points)

Preferred block size for 1:1:2 block ratio:
1+1+2 = 4*2 = 8 is the ideal block size.
Generating a randomization schedule for 9600 prospective subjects in 10 study centers, three treatments, 1:1:2 allocation, block size of 8.
10*X*8 = 9600 = 120; there should be 120 blocks of size 8 at 10 study centers for a prospective 9600 participants.

SAS Code:

```
proc plan seed = 10;
        factors site = 10 ordered blocks = 120 ordered trt= 8 random/noprint;
        output out = rsched trt cvals = ('A' 'A' 'B' 'B' 'C' 'C' 'C' 'C');
run;quit;
data rsched2;
        set rsched;
        by site;
        if first.site then subject=0; subject+1;
        if site=1 then subid=subject+100;
        if site=2 then subid=subject+200;
        if site=3 then subid=subject+300;
        if site=4 then subid=subject+400;
        if site=5 then subid=subject+500;
        if site=6 then subid=subject+600;
        if site=7 then subid=subject+700;
        if site=8 then subid=subject+800;
        if site=9 then subid=subject+900;
        if site=10 then subid=subject+1000;
run;
proc print data=rsched2(obs=20) noobs;
by site;
var subid trt;
run;
```

Subject ID	Treatment Assignment
101	C
102	C
103	C
104	A
105	C
106	B
107	A
108	B
109	B
110	C
111	A

112	C
113	C
114	C
115	A
116	B
117	C
118	C
119	C
120	C

Table 1. SAS output of randomization schedule for the first country (site = 1, observations = 20)

PART 2: TABLE OF BASELINE COMPARABILITY

2. Create a table to assess baseline comparability of treatment groups including the following baseline variables: SEX, AGE, RSBP, STYPE. Your table should have one column for each treatment group, the mean and SD for continuous variables and N (%) for categorical variables. Do the treatment group look well balanced on these factors? (*You do not need to include any p-values in this table.)* (5 points)

	High-Dose (N = 2307)	**Low-Dose (N = 2432)**	**Placebo (N = 4858)**
Sex (N, %)			
Female	1092 (47.33)	1115 (45.85)	2309 (47.53)
Male	1215 (52.67)	1317 (54.15)	2549 (52.47)
Age (mean, 95% CI)	71.99 (71.52, 72.47)	71.50 (71.04, 71.96)	71.77 (71.44, 72.10)
SBP in mmHg (mean, 95% CI)	159.40 (158.28, 160.52)	160.97 (159.84, 162.10)	159.74 (158.97, 160.51)
Stroke Subtype (N, %)			
LACS	560 (24.27)	606 (24.92)	1134 (23.34)
OTH	8 (0.35)	4 (0.16)	12 (0.25)
PACS	954 (41.35)	934 (38.40)	1982 (40.80)
POCS	262 (11.36)	313 (12.87)	529 (10.89)
TACS	523 (22.67)	575 (23.64)	1201 (24.72)

Table 2. Baseline characteristics of sex, age, SBP, and stroke subtype stratified by treatment group

Overall, baseline characteristics in treatment groups look well-balanced on sex, age, SBP, and stroke subtype. The first slight difference (~2% of group) observed is in the sex distribution of low-dose compared to high-dose and placebo. Age and SBP are similar values across all three treatment groups. Stroke subtype also has slight differences across groups, especially in subtype PACS and TACS. These differences aren't drastic enough to be concerning. The OTH group has few observations across all three treatment groups which may lead to some issues in later analysis.

PART 3: ANALYSIS OF THE PRIMARY OUTCOME

3. The study is considered a success if at least one of the treatment doses has a significantly lower rate of recurrent stroke within 14 days compared to the placebo group. With this in mind, please state the two-sided null and alternative hypotheses of interest. (5 points)

H_0: $p_{\text{high-dose treatment}} = p_{\text{placebo}}$ or $p_{\text{low-dose treatment}} = p_{\text{placebo}}$
H_A: $p_{\text{high-dose treatment}} \; != p_{\text{placebo}}$ or $p_{\text{low-dose treatment}} \; != p_{\text{placebo}}$

The null hypothesis is that the proportion of recurrent stoke events within 14 days is the same in the high-dose treatment group and the placebo group, or that the proportion of recurrent stoke events within 14 days is the same in the low-dose treatment group and the placebo group. The alternative hypothesis is

that the proportion of recurrent stoke events within 14 days is different in the high-dose treatment group and the placebo group, or that the proportion of recurrent stoke events within 14 days is different in the low-dose treatment group and the placebo group. This will be a two-sided test using a FWER of 0.05 and it will be considered a success if at least one of the two observed proportions in the two treatment groups is lower than in the placebo group.

4. Provide descriptive statistics (N, %) for recurrent stroke within 14 days for each treatment group. No formal table, p-values or confidence intervals are needed at this point (SAS output including descriptive statistics is fine). Do you think the study will be a "success" based on these preliminary results? Please explain briefly. (5 points)

	High-Dose (N = 2307)	**Low-Dose (N = 2432)**	**Placebo (N = 4858)**
Stroke (N, %)			
No	2213 (95.93)	2371 (97.49)	4687 (96.48)
Yes	94 (4.07)	61 (2.51)	171 (3.52)

Table 3. Descriptive statistics for recurrent stroke within 14 days, stratified by treatment group

Based off the table alone, I think this study will be a success with the outcome that low-dose reduces the incidence of recurrent stroke within 14 days compared to the placebo group. I do not think the high-dose comparison will result in rejecting the specified null hypothesis, but the study will be a success from one group showing a significant difference.

5. Perform the appropriate formal statistical treatment group comparisons on the rate of the primary endpoint of recurrent stroke within 14 days (i.e., test your null hypotheses in question 3), controlling the overall experiment-wise error rate at a 0.05 level of significance. Please do not adjust for any covariates. When writing your results, state the statistical approach you use to control the experiment-wise error rate. State your null and alternative hypotheses of interest, your p-values, and which null hypothesis(es) that you reject. Provide appropriate pairwise confidence intervals of the difference between treatment rates, again controlling the experiment-wise error rate at a 0.05 level of significance. (10 points)

To test the following hypotheses, we can use a chi-square test of high-dose and low-dose treatments to placebo treatment. Since we have three treatment arms, we will be performing two chi-square tests and need to correct the p-values using a Bonferroni adjustment to control the experiment-wise error rate at 0.05. Expected cell counts are above five, so we can use a chi-square test to determine if the difference of proportions between groups is significantly significant.

At the alpha = 0.05 level of significance, we test the following:
H_0: $p_{\text{high-dose treatment}} = p_{\text{placebo}}$ or $p_{\text{low-dose treatment}} = p_{\text{placebo}}$
H_A: $p_{\text{high-dose treatment}} \mathrel{!}= p_{\text{placebo}}$ or $p_{\text{low-dose treatment}} \mathrel{!}= p_{\text{placebo}}$

This is a two-sided test, with success when the proportion(s) of the experimental treatment groups high-dose and low-dose are lower than the placebo group.

Treatment	**Chi-Square**	**P (Corrected)**	**Risk Difference (95% CI)**
Low Dose v. Placebo	5.38	0.02*2 = 0.04	-0.01 (-0.019, -0.001)
High Dose v. Placebo	1.35	0.25*2 = 0.5	0.006 (-0.005, 0.017)

Table 4. Results testing global significance and individual treatment difference in each treatment group, with placebo as reference

Since the chi-square value is significantly large for the low dose versus placebo group (5.38, p = 0.04 < 0.05), we can reject the null hypothesis and conclude that there is a difference in the proportions of individuals with a recurrent stoke event within 14 days between the low dose group and the placebo

group. The risk for recurrent stroke in the low dose group is 1% lower than the placebo group (95% CI: -0.02; -0.001), and since the adjusted CI does not contain 0, this finding is significant.

Since the chi-square value is not significantly large for the high dose versus placebo group (Chi-Square = 1.35, p = 0.5 > 0.05), we cannot reject the null hypothesis and conclude that there is not a difference in the proportions of individuals with a recurrent stoke event within 14 days between the high dose group and the placebo group. The risk for recurrent stroke in the high dose group is 0.6% greater than the placebo group (95% CI: -0.01; 0.02) and since the adjusted CI contains 0, this difference is not significant when controlling for a FWER of 0.05.

6. Repeat question 5, now adjusting for SEX. Compare the effect of treatment in this model to the model in question 5. Is the treatment effect similar to the unadjusted results in question 5? Discuss why there might or might not be a difference between the adjusted and unadjusted models. (10 points)

To test the following hypotheses, we can use a chi-square test of high-dose and low-dose treatments to placebo treatment. Since we have three treatment arms, we will be performing two chi-square tests and need to correct the p-values using a Bonferroni adjustment to control the experiment-wise error rate at 0.05. Expected cell counts are above five, so we can use a chi-square test to determine if the difference of proportions between groups is significantly significant.

At the alpha = 0.05 level of significance, we test the following:
H_0: $p_{\text{high-dose treatment}} = p_{\text{placebo}}$ or $p_{\text{low-dose treatment}} = p_{\text{placebo}}$
H_A: $p_{\text{high-dose treatment}} \text{ != } p_{\text{placebo}}$ or $p_{\text{low-dose treatment}} \text{ != } p_{\text{placebo}}$

This is a two-sided test, with success when the proportion(s) of the experimental treatment groups high-dose and low-dose are lower than the placebo group.

Global Wald Chi-Square for Treatment: 9.1280, df = 2, p = 0.01
At least one of the three treatments is different from the other.

Logit(probability of recurrent stroke) = B0 + B1(high dose) + B2(low dose) + B3(Female Sex)
= -3.36 + 0.15(HD) – 0.35(LD) + 0.09(F)

Treatment	**Chi-Square**	**OR (95% CI)**	**P (Corrected)**
Low Dose v. Placebo	5.29	0.71 (0.53, 0.95)	0.04
High Dose v. Placebo	1.35	1.16 (0.90, 1.51)	0.49

Table 5. Results of logistic regression testing global significance and individual treatment difference in each treatment group, with Placebo as reference and adjusting for sex.

Since the chi-square value is significantly large for the low dose versus placebo group (5.29, p = 0.04 < 0.05), we can reject the null hypothesis and conclude that there is a difference in the proportions of individuals with a recurrent stoke event within 14 days between the low dose group and the placebo group adjusting for sex.
The odds of recurrent stroke event are 29% lower for those in the low dose treatment group compared to the placebo. This finding is significant since the confidence interval does not contain the null value of 1 (OR = 0.71, 95% CI: 0.53; 0.95).

Since the chi-square value is not significantly large for the high dose versus placebo group (Chi-Square = 1.35, p = 0.49 > 0.05), we cannot reject the null hypothesis and conclude that there is not a difference in the proportions of individuals with a recurrent stoke event within 14 days between the high dose group and the placebo group adjusting for sex.

The odds of recurrent stroke event are 16% higher for those in the high dose treatment group compared to the placebo. This finding is not significant since the confidence interval contains the null value of 1 (OR = 1.16, 95% CI: 0.90; 1.51).

The treatment effect is similar between the adjusted and unadjusted models. This would indicate that sex is not a confounder, and it does not change the effect of the treatment on the incidence of recurrent stroke event.

7. In the model from part 6, does SEX appear to be an important predictor of STROKE14? Why or why not? Can you tell from this model whether the effect of the treatment is different for females and males? Why or why not? (10 points)

The effect estimation of being a female in the study is not significant at the alpha = 0.05 level of significance (regression coefficient for female = 0.09, p = 0.41). Female participants have 1.10 times the odds of recurrent stroke within 14 days compared to male participants keeping treatment group constant (95% CI: 0.88; 1.37). Since this interval contains the null value of 1 there is not a significant difference between female and male sex groups.

8. The above analyses included participants who were non-compliance so that it reflects an intention-to-treat (ITT) approach. Now perform a per-protocol (PP) type analysis by only including study subjects who were compliant to both aspirin (CMPLASP) and heparin (CMPLHEP) and redoing the analysis in part 5. How does the per-protocol analysis compare to the intention-to-treat analysis performed in part 5? Do you think either analysis could be biased? Explain your answer. (10 points)

To test the following hypotheses in the compliant subgroup, we can use a chi-square test of high-dose and low-dose treatments to placebo treatment. Since we are performing a global test, we can use the alpha level of 0.05.

At the alpha = 0.05 level of significance, we test the following:
H_0: $p_{\text{high-dose treatment}} = p_{\text{placebo}}$ or $p_{\text{low-dose treatment}} = p_{\text{placebo}}$
H_A: $p_{\text{high-dose treatment}} \mathrel{!}= p_{\text{placebo}}$ or $p_{\text{low-dose treatment}} \mathrel{!}= p_{\text{placebo}}$

This is a two-sided test, with success when the proportion(s) of the experimental treatment groups high-dose and low-dose are lower than the placebo group. To control for FWER at 0.05, we will use a Bonferroni correction.

Treatment (PP)	**Chi-Square**	**P (Corrected)**	**Risk Difference (95% CI)**
Low Dose v. Placebo	6.45	0.01*2 = 0.02	-0.01 (-0.019, -0.002)
High Dose v. Placebo	1.52	0.22*2 = 0.44	0.006 (-0.005, 0.017)

Chi-Square Score	**P**
10.71	0.005
Treatment	**OR Estimate (95% CI)**
High Dose v. Placebo	1.20 (0.86, 1.68)
Low Dose v. Placebo	0.64 (0.43, 0.95)

Table 6. Results of crude logistic regression testing global significance and individual treatment difference in each treatment group using compliant individuals, with placebo as reference

Since the chi-square value is significantly large (10.71, p = 0.01 < 0.05, 2 df), we can conclude that at least one estimate of the effect of treatment group when referencing placebo group is not 0.

Participants randomly assigned to the high-dose group had 1.20 times the odds of recurrent stroke incidence within 14 days compared to those in the placebo group in the compliant individual subgroup (95% CI: 0.86; 1.68). Since this interval contains the null value of 1, we do not have enough evidence to reject the null hypothesis that there is a difference in the occurrence of recurrent stroke events in the high dose and placebo groups.

Participants randomly assigned to the low dose group had 0.64 times the odds of a recurrent stroke event within 14 days compared to those in the placebo group (95% CI: 0.43, 0.95). This value is significant since the interval does not contain 1. We have statistically significant evidence to reject the null hypothesis and can conclude that the low dose group has a lower recurrent stroke event proportion than the placebo group since the OR is below 1 (about a 36% reduction in odds).

The per protocol analysis and the intention to treat analysis both capture the same directions of association for high dose associations and low dose associations. We found that the low dose comparison is more significant in the PP compared to ITT. The PP is likely to be biased - individuals in a sample who are completely complaint to interventions are not representative of what we would see in a real-world population, therefore we likely would not observe in a real population what we observed in the PP analysis. We are more likely to observe what we see in the ITT analysis in a population, so the effects we observe from that analysis represent what we could see when the dosages are available to the public.

PART 4: ANALYSIS OF SECONDARY OUTCOMES

9. The secondary endpoint of this study is death within 14 days. This outcome will be treated as a time to event outcome with variable TD representing the time to death or censoring in days. It is hoped that at least one dose of Heparin has a lower hazard of death than the placebo. Please state the two-sided null and alternative hypotheses of interest for this secondary endpoint. (5 points)

H0: $HR_{low} = 1$; there is no difference in hazard of death between low dose heparin and placebo
or $HR_{high} = 1$; there is no difference in hazard of death between high dose heparin and placebo
HA: there is a difference in hazard of death between low dose of heparin and placebo treatment groups
or there is a difference in hazard of death between high dose of heparin and placebo treatment groups

10. Estimate the survival probabilities for each treatment group using PROC LIFETEST. What are the overall survival probabilities in the three group? (*note that there will be a lot of output to look through!*) Include the survival plot in your report. Carefully inspect and describe the survival curves for the three treatment groups. Do you think there are important differences between groups in survival during the first 14 days of follow-up? Please explain briefly without carrying out any statistical tests. (10 points)

H0: there is no difference in survival between high dose of Heparin, low dose of Heparin, and placebo treatment groups.
HA: there is a difference in survival between high dose of Heparin, low dose of Heparin, and placebo treatment groups

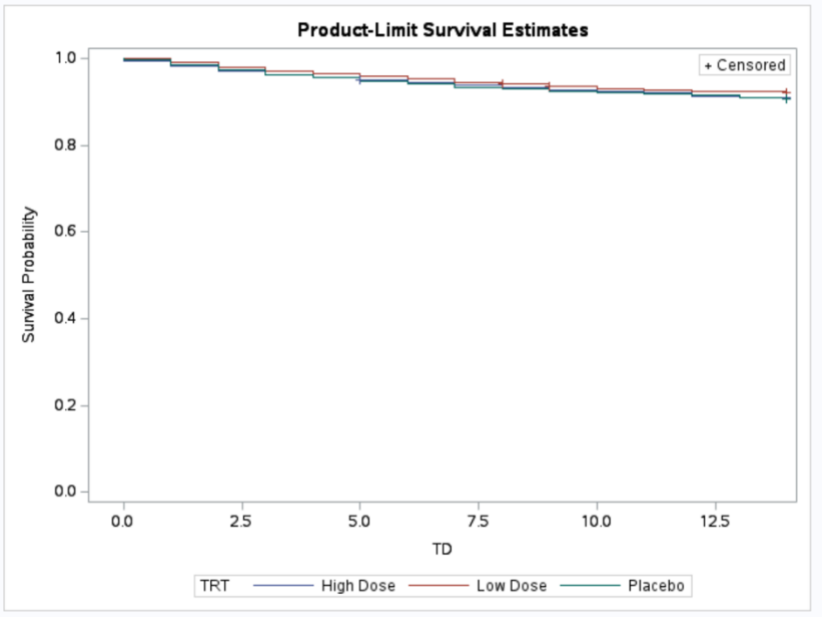
	High Dose	**Low Dose**	**Placebo**
Time Interval (Days)	**KM Survival Probability**	**KM Survival Probability**	**KM Survival Probability**
0	0.9952	0.9988	0.9957
1	0.9809	0.9901	0.9850
2	0.9710	0.9794	0.9749
3	0.9606	0.9716	0.9632
4	0.9571	0.9659	0.9559
5	0.9497	0.9576	0.9477
6	0.9432	0.9519	0.9401

7	0.9384	0.9457	0.9335
8	0.9315	0.9412	0.9294
9	0.9276	0.9346	0.9249
10	0.9233	0.9309	0.9214
11	0.9202	0.9284	0.9168
12	0.9137	0.9247	0.9144
13	0.9107	0.9227	0.9103
14	0.9081	0.9202	0.9070

Table 7. KM Survival Probability estimates for each treatment group per day

After 14 days, the overall survival probability for the high dose treatment group is 90.81%, 92.02% for the low dose treatment group, and 90.70% for the placebo group.

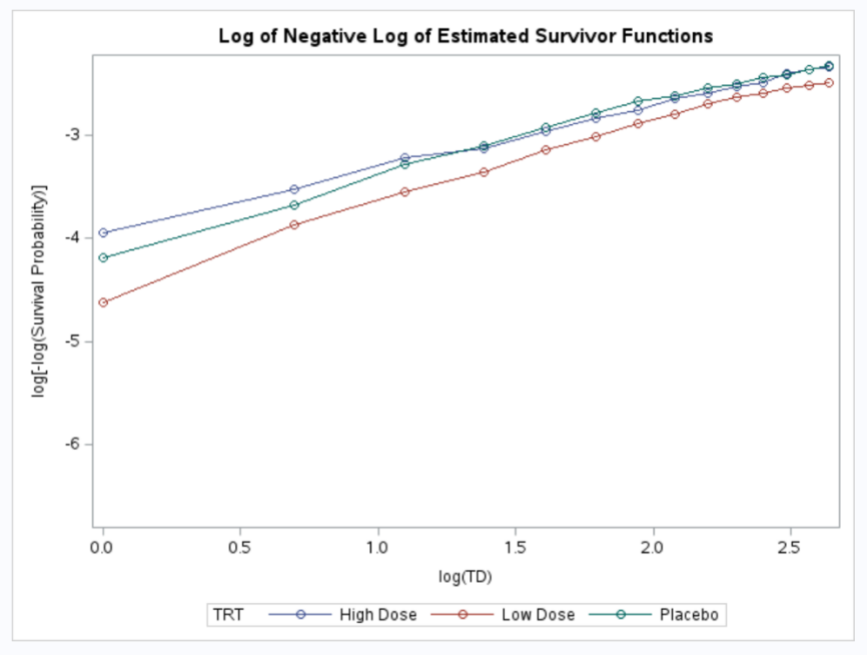
Figure 1. KM Curve for high dose, low dose, and placebo treatment groups



All curves in the plot appear to follow the same trend – decreasing the most from 0 to 5 days, then gradually decreasing from 5-14 days, flattening out around 10 days. The low dose group appears to have a slightly higher survival probability compared to high dose and placebo treatment groups. High dose and placebo groups appear to overlap one another on most of the 14 days plotted. I do not believe this differences in survival in the first 14 days of follow-up will be significant since there are very slight variations from one group to another.

11. Use PROC LIFETEST to assess whether the proportional hazards assumption for the treatment effect is met. Include any relevant output to justify your response. (5 points)

Figure 2. Log-log survival plot of high dose, low dose, and placebo groups



It does not appear that the proportional hazards assumption for treatment effect is met. High dose and placebo overlap and cross, indicating a violation in proportional hazards.

12. Use a Bonferroni correction procedure to test the null and alternative hypotheses in question 9, controlling the experiment-wise error rate at 0.05 across the hypotheses. Present hazard ratios of death within 14 days of each treatment dose vs. placebo and their appropriate multiple-comparison-adjusted confidence intervals. (10 points)

H0: $HR_{low} = 1$; there is no difference in hazard of death between low dose heparin and placebo
or $HR_{high} = 1$; there is no difference in hazard of death between high dose heparin and placebo
HA: there is a difference in hazard of death between low dose of heparin and placebo treatment groups
or there is a difference in hazard of death between high dose of heparin and placebo treatment groups

Since we will be testing two hypotheses, we will use a Bonferroni correction for two tests to control for an experiment-wise error rate of 0.05.

	HR (95% Corrected CI)	**Z Statistic (P)**
High Dose	0.99 (0.82; 1.19)	-0.15 (1.00)
Low Dose	0.85 (0.70; 1.03)	-1.88 (0.12)

Table 8. High dose and low dose hazards compared to placebo group on the secondary outcome of death

The high dose group has 0.99 times the hazard of death within 14 days of treatment compared to the placebo group (95% CI 0.82, 1.19). Since this interval contains the null value of 1, the finding is not significant.
The low dose group has 0.85 times the hazard of death within 14 days of treatment compared to the placebo group (95% CI 0.70, 1.03). Since this interval contains the null value of 1, the finding is not significant.
We do not have significant evidence to reject the null in the comparison between high dose and placebo and low dose and placebo, and state that there is a difference in hazard in death between at least one dose of heparin and placebo.

PART 5: CONCLUSION

13. Given all the results you compiled above for the primary and secondary outcomes, is the study "successful"? Do you think that one or both of the doses of Heparin should be approved by the FDA based on this study? Please explain briefly. (10 points)

The primary outcome of the study was that at least one of the treatment doses has a significantly lower rate of recurrent stroke within 14 days compared to the placebo group, and the study was successful with this outcome. Low dose of heparin was significantly associated with lower odds of recurrent stroke within 14 days when compared to the placebo group (adjusted OR by sex = 0.71, 95% CI: 0.53; 0.95). The low dose should be approved by the FDA based on this study. High dose was not found to be significantly associated with recurrent stroke within 14 days (adjusted OR by sex = 1.16, 95% CI: 0.90; 1.51). The high dose should not be recommended for approval by the FDA.
We did not find low dose or high dose to be significantly associated with death after 14 days when compared to the placebo (low dose HR = 0.85, 95% CI: 0.70; 1.03 and high dose HR = 0.99, 95% CI: 0.82; 1.19). This investigation was not the primary goal of the analysis, so the low dose of heparin should still be approved by the FDA even if not found significant in reducing death.

SAS Code and Output is attached as a separate file – survival probabilities over 14 days were omitted from output for ease of review.