

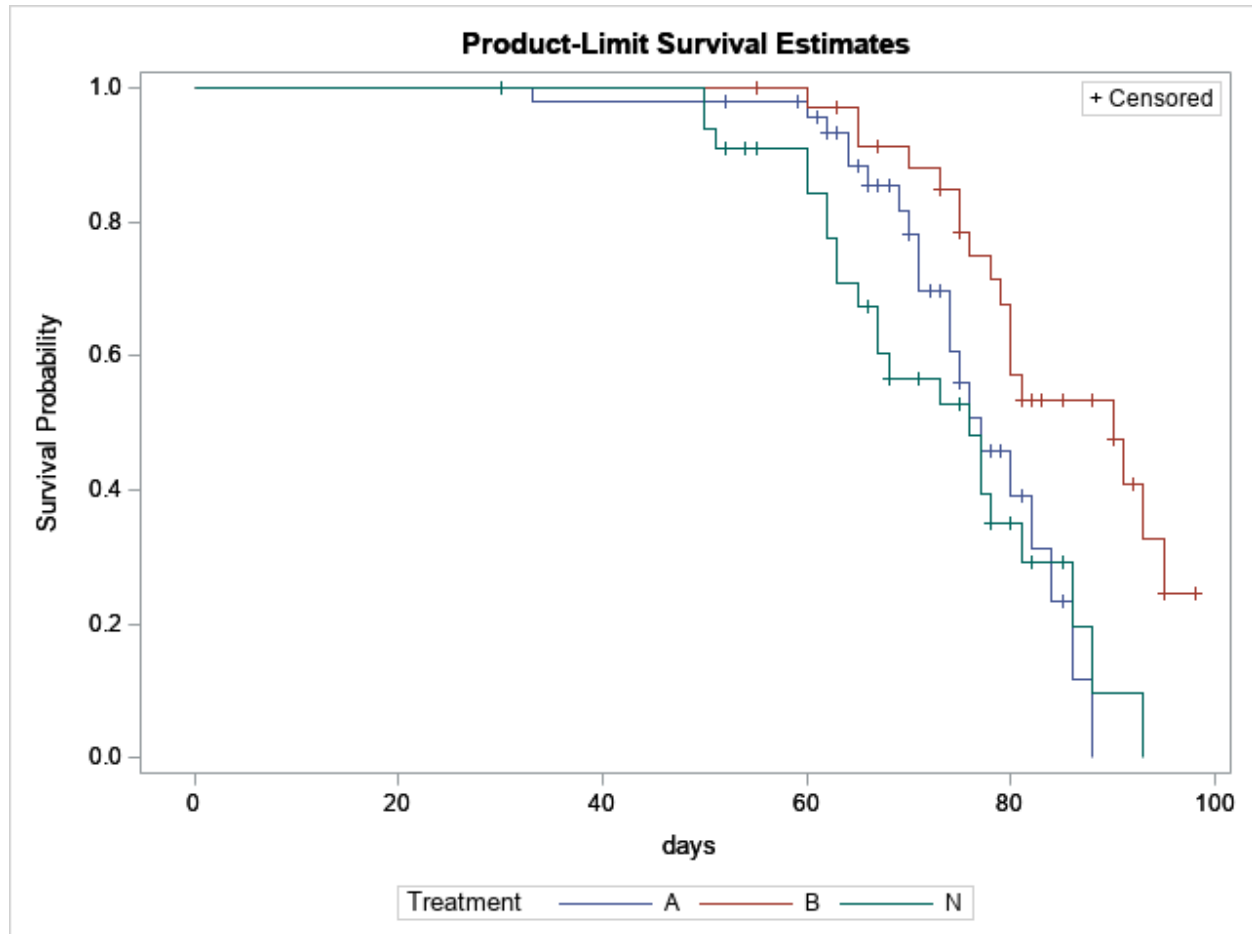
A phase III superiority study was conducted to determine if a new treatment for depression (treatment N) promotes remission better than *two* available standard of care therapies (treatment A and B). There is no placebo in this study due to the severity of the disease under study.

The study included 117 subjects with a recent diagnosis of depression were enrolled and treated in this study. The primary outcome was time to depression remission (9-item Patient Health Questionnaire [PHQ-9] score <5 and persistent depressive symptoms (PDSs; PHQ-9 score ≥ 10). All patients were to be followed for a fixed time of 100 days, or until remission or loss to follow-up or death occurred. Note that here, the outcome is 1="remission" (a positive outcome). Thus, "survival" actually means no remission and "failure" means depression remission; please keep this in mind as you interpret your results.

The variables in the dataset are:

- REMISS indicates Remission Status (0=No,1=Yes)
- DEATH indicates Death (1=Died, 0=Lost to follow-up or had remission)
- DAYS indicates days to remission (if REMISS=1) or days until no longer in study (if REMISS=0 due to loss to follow-up or death)
- TRT indicates Treatment Group (N, A, B)
- AGEGRP indicates Age Group (Young or Old)

Present a Kaplan-Meier survival plot for N vs. A vs. B (all 3 treatments should be on the same graph). Describe the nature of the survival experience in the 3 groups including timing of events and loss to follow-up, without any formal p-value. Which treatment appears to be best in promoting remission (15 points)?



The treatment A group had the first event at 33 days, and then events steadily occur starting at 64 days until the end of the study. Subjects in the treatment B group were most resistant to depression remission; the first event didn't occur until 60 days. The treatment N group didn't have an event until 50 days, but a larger proportion of subjects had depression remission from that point on. Treatment N seems to be the best in promoting depression remission, while Treatment B appears to be the worst in doing so.

What proportion of patients experienced depression remission by 70 days in each of the 3 groups (10 points)? (Remember that remission is a positive outcome: 'Survival' indicates that a patient did not experience remission.)

Treatment	Depression Remission at 70 days
A	21.97%
B	11.89%
N	56.71%

Report median time to remission ('median survival') for the three groups and discuss which treatment appears to be more effective (again, do not base your conclusions on p-values, this is subjective)

Treatment	Median Time to Remission
A	77 days
B	90 days
N	76 days

Treatment N appears to be the most effective because it has the least time to depression remission.

Present the null hypothesis and p-value of separate pairwise log-rank tests of N vs. A and N vs. B. Do the results agree with your conclusion in part a (10 points)?

Note: Please ignore the issue of multiple testing due to multiple comparisons, we will cover that in a few weeks.

$H_0: S_N(t) = S_A(t)$ The survival distribution for treatments N and A are the same.

$H_1: S_N(t) = (S_A(t))^\theta$ The survival distribution for one group is a power of the other.

$H_0: S_N(t) = S_B(t)$ The survival distribution for treatments N and B are the same.

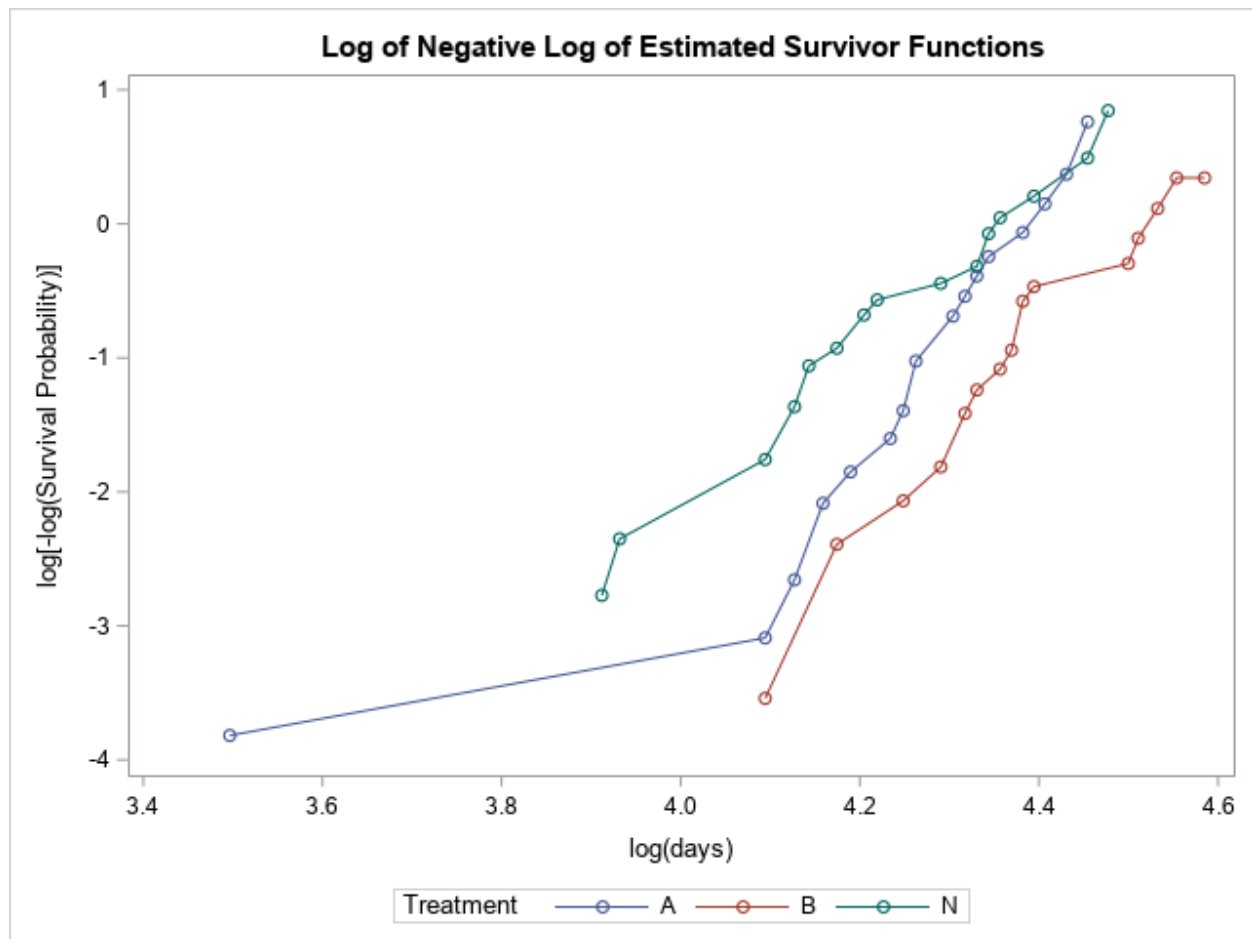
$H_1: S_N(t) = (S_B(t))^\theta$ The survival distribution for one group is a power of the other.

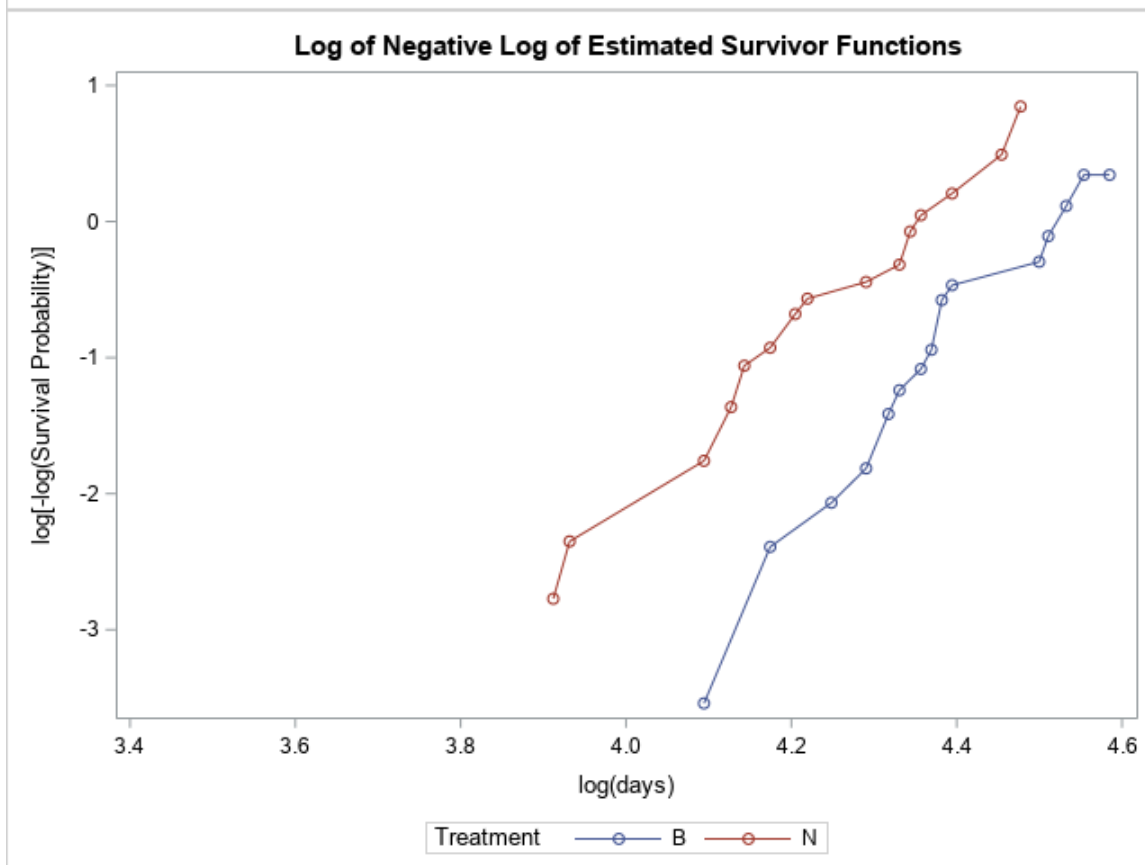
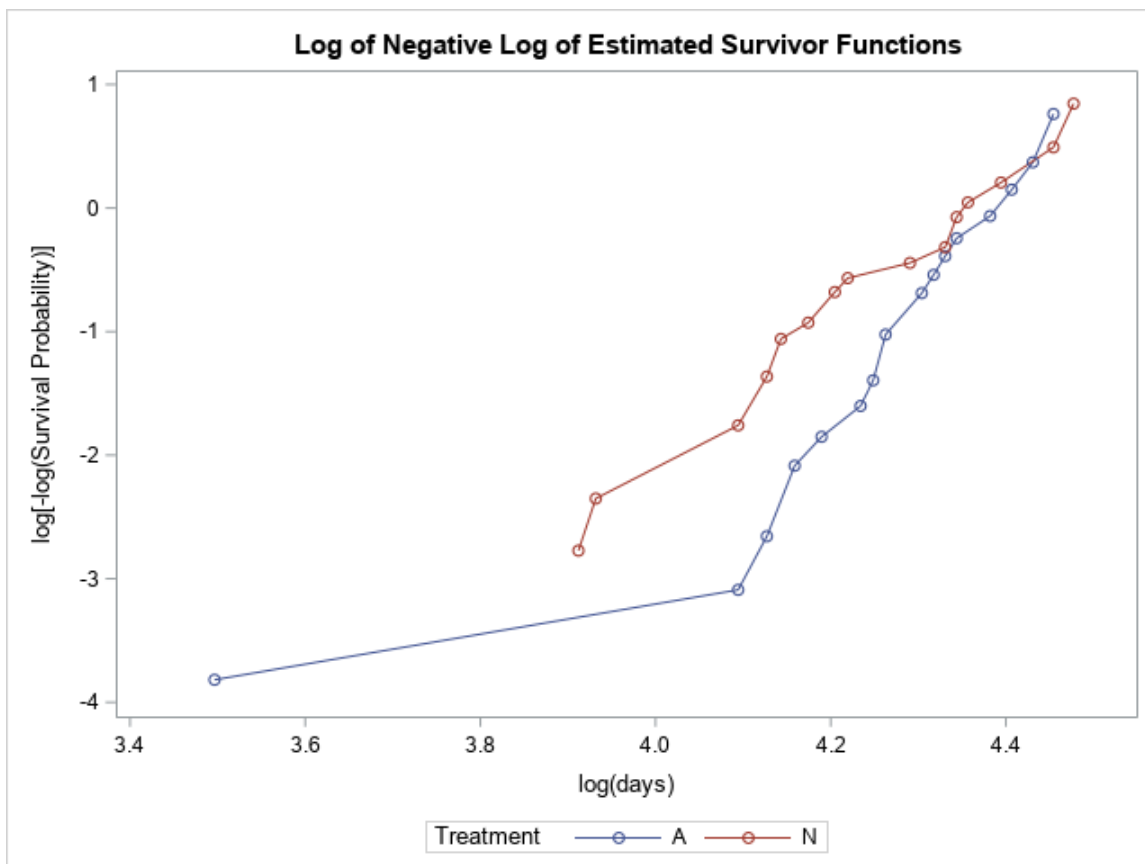
A log-rank test was used to test whether the survival distributions for treatment groups N and A were different. The chi-squared statistic was 0.6112 and the resulting p-value was 0.4343. With a p-value greater than the $\alpha=0.05$ significance level, the null hypothesis of there being a difference in the survival distributions between treatment groups N and A was not rejected. There is insufficient evidence to conclude that treatment N is more effective for depression remission than treatment A.

A log-rank test was used to test whether the survival distributions for treatment groups N and B were different. The chi-squared statistic was 10.4435 and the resulting p-value was 0.0012. With a p-value less than the $\alpha=0.05$ significance level, the null hypothesis of there being a difference in the survival distributions between treatment groups N and B was rejected. There is evidence suggesting that treatment N is more effective for depression remission than treatment B.

Present the log-log survival plot. Does the proportional hazards assumption appear to be met? Regardless of the answer to this question, continue with the remainder of the assignment, even if you would not in practice.

Based on the log-log survival plots, the proportional hazards assumption is met between N and B, but not between N and A because the curves cross a couple of times.





Assume that regulatory approval requires the new treatment to be more effective than both standard of care treatments A and B. Use one regression model to determine if the FDA would approve the new treatment. Present and interpret the hazard ratio and p-value for treatment A vs N and the hazard ratio and p-value for treatment B vs N.

$H_0: HR = 1, \beta = 1$

The hazard ratios for all treatments are the same.

$H_1: HR \neq 1, \beta \neq 1$

The hazard ratios for all the treatments are not the same.

A Cox proportional hazards regression analysis was used to test whether the hazard ratios for all the treatments were the same. Comparing treatments N and A, the chi-squared statistic was 0.6433 and the resulting p-value was 0.4225. With a p-value greater than the $\alpha=0.05$ significance level, the null hypothesis of there being a difference in hazard ratios of treatments N and A was not rejected. The hazard of depression remission for treatment A is 0.778 of the hazard for treatment N (95% confidence interval: 0.421-1.437), but there is insufficient evidence to conclude that treatment N is more effective than treatment A.

Comparing treatments N and B, the chi-squared statistic was 10.7099 and the resulting p-value was 0.0011. With a p-value less than the $\alpha=0.05$ significance level, the null hypothesis of there being a difference in hazard ratios of treatments N and B was rejected. There is evidence suggesting that the hazard of depression remission for treatment B is 0.338 of the hazard for treatment N (95% confidence interval: 0.177-0.647), making treatment N more effective at bringing about depression remission.

Could differential competing risk due to death in the three treatment groups be biasing our results? Report the proportion who died in each treatment group, the appropriate statistical test to compare these proportions and its associated p-value, and your conclusions.

Treatment	Proportion Death
A	28.26%
B	27.03%
N	26.47%

A chi-squared test was used to test whether there was a difference in proportion of death between the three treatment groups. The chi-squared statistic was 0.0344 with 2 degrees of freedom and resulting p-value was 0.9830. With a p-value greater than the $\alpha=0.05$ significance level, the null hypothesis of there being no difference in proportion of death between the treatment groups was not rejected. There is insufficient evidence to conclude that the competing risk of death was different between the three treatment groups.

Write up the methodology and results of your Cox regression analyses, including confidence intervals.

A Cox proportional hazards regression analysis was used to test whether the hazard ratios for all the treatments were the same. Comparing treatments N and A, the chi-squared statistic was 0.6 and the resulting p-value was 0.423. With a p-value greater than the $\alpha=0.05$ significance level, the null hypothesis of there being a difference in hazard ratios of treatments N and A was not rejected. The hazard of depression remission for treatment A is 77.8% of the hazard for treatment N (95% confidence interval: 42.1%-143.7%), but there is insufficient evidence to conclude that treatment N is more effective than treatment A.

Comparing treatments N and B, the chi-squared statistic was 10.7 and the resulting p-value was 0.001. With a p-value less than the $\alpha=0.05$ significance level, the null hypothesis of there being a difference in hazard ratios of treatments N and B was rejected. There is evidence suggesting that the hazard of depression remission for treatment B is 33.8% of the hazard for treatment N (95% confidence interval: 17.7%-64.7%), making treatment N more effective at bringing about depression remission.