## Interim Analyses

A randomized trial for Hepatitis C is planned with 1:1 randomization to a new treatment versus a standard of care treatment (control). The initial goal is to enroll 1250 patients in both treatments combined. The primary outcome is sustained virological response ("SVR") at 6 months after randomization (meaning hepatitis C virus is not detected in the blood). The investigators plan to carry out 4 interim analyses (***plus*** the final analysis) and will consider stopping the study at an interim analysis if there is a significant difference in SVR at 6 months between the two treatments using O'Brien-Fleming two-sided significance levels. Suppose we are at the design stage of the study, and we are *planning* on enrolling 1250 patients. Suppose that the first interim analysis is *scheduled* after the first 500 patients (250 per group) are treated and followed for 6 months, the second interim analysis is *scheduled* after the first 700 patients (350 per group) are treated and followed for 6 months, the third interim analysis is *scheduled* after the first 900 patients (450 per group) are treated and followed for 6 months, the fourth interim analysis is *scheduled* after the first 1100 (550 per group) are treated and followed for 6 months, and the final analysis is *planned* after all 1250 patients have been treated and followed.

a. Find the O'Brien-Fleming two-sided significance levels for the 5 analyses (the 4 interim analyses, plus the final). Use SAS PROC SEQDESIGN, and use it at an overall two-sided 0.05 level of significance. In a nicely formatted table, please provide the alpha spent, critical values, significance levels at each interim stage and the final stage. Explain *briefly* the general difference between a significance level (at a given interim analysis) versus "alpha-spent" (at a given interim analysis).

| **Analysis** | **k** | **n (sample at time look k)** | **s (proportion of information/sample accrued)** | **Alpha spent** | **Critical value** | **Significance level** |
|---|---|---|---|---|---|---|
| Interim 1 | 1 | 500 | 500/1250 = 0.40 | 0.00039*2 = 0.000784 | 3.35687 | 0.000392*2 = 0.0007884 |
| Interim 2 | 2 | 700 | 700/1250 = 0.56 | 0.00274*2 = 0.00548 | 2.79258 | 0.00261*2 = 0.00522 |
| Interim 3 | 3 | 900 | 900/1250 = 0.72 | 0.00825*2 = 0.0165 | 2.43797 | 0.00739*2 = 0.01478 |
| Interim 4 | 4 | 1100 | 1100/1250 = 0.88 | 0.01688*2 = 0.03376 | 2.19077 | 0.01423*2 = 0.02846 |
| Final | 5 | 1250 | 1250/1250 = 1 | 0.02500*2 = 0.05 | 2.06491 | 0.01947*2 = 0.03894 |

The general difference between a significance level (at a given interim analysis) and "alpha-spent" (at a given interim analysis) is that the significance level represents the maximum p-value that can be considered statistically significant at that interim analysis, while the alpha-spent represents the amount of the overall alpha level that has already been used up by previous analyses. In other words, the significance level is a fixed threshold for rejecting the null hypothesis at a given interim analysis, while the alpha-spent is a cumulative measure of how much evidence has been accumulated so far for or against the null hypothesis.

b. It's now time for the first interim analysis comparing treatments. It turns out there are 450 total patients in this first analysis (225 per group) instead of the planned 500. Revise the O'Brien-Fleming two-sided significance levels to use for this first analysis and subsequent analyses 2 through 5, still assuming a planned total of 1250 patients and still planning for the remaining interim analyses to be carried out at the planned sample sizes discussed above (700, 900, and 1100). In a nicely formatted table, please provide the alpha spent, critical values, significance levels at each interim stage and the final stage.

| **Analysis** | **k** | **n (sample at time look k)** | **s (proportion of information/sample accrued)** | **Alpha spent** | **Critical value** | **Significance level** |
|---|---|---|---|---|---|---|
| Interim 1 | 1 | 450 | 450/1250 = 0.36 | 0.00019*2 = 0.00038 | 3.55748 | 0.0001872*2 = 0.0003744 |
| Interim 2 | 2 | 700 | 700/1250 = 0.56 | 0.00274*2 = 0.00548 | 2.78462 | 0.00268*2 = 0.00536 |
| Interim 3 | 3 | 900 | 900/1250 = 0.72 | 0.00825*2 = 0.0165 | 2.43726 | 0.00740*2 = 0.0148 |
| Interim 4 | 4 | 1100 | 1100/1250 = 0.88 | 0.01688*2 = 0.03376 | 2.19057 | 0.01424*2 = 0.02848 |
| Final | 5 | 1250 | 1250/1250 = 1 | 0.02500*2 = 0.05 | 2.06481 | 0.01947*2 = 0.03894 |

The main difference in these revised levels compared to the original levels is the reduced spending function for the first interim analysis due to the smaller than planned sample size. This results in a slightly lower alpha-spent and a slightly more conservative critical value and significance level for the first interim analysis. The subsequent interim analyses are not affected by the sample size discrepancy and therefore have the same

planned spending function as before.

c) After the first but before the second interim analysis, the authors decided to increase the total sample size for the Primary Efficacy Population from 1250 to 1750 (875 per group). Re-calculate the two-sided significance levels to use for analyses 2 through 5, keeping in mind that one interim analysis has already been done on 450 patients at a given significance level, and that this first significance level and the corresponding alpha that was spent cannot be changed. Also, when calculating the revised significance level values for the remaining analyses, please do it under these assumptions:

(1) Despite the increase in sample size, the analyses will still be carried out on the first 700, 900, and 1100 total patients, respectively, but that the final analysis will, again, be conducted on the 1750 total patients.

(2) At the time of each interim analysis, we would like to only spend the same amount of alpha that we spent at each interim analysis in Part B (e.g., the alpha-spent by the second interim analysis under this new total sample size of 1750 should match the alpha-spent by the second interim analysis under the old total sample size of 1250 in Part B, and similarly for analyses 3-5).

| **Analysis** | **k** | **n (sample at time look k)** | **s (proportion of information/sample accrued)** | **Alpha spent** | **Critical value** | **Significance level** |
|---|---|---|---|---|---|---|
| Interim 2 | 2 | 700 | 700/1750 = 0.40 | 0.00274*2 = 0.00548 | 2.78511 | 0.00268*2 = 0.00536 |
| Interim 3 | 3 | 900 | 900/1750 = 0.514 | 0.00825*2 = 0.0165 | 2.43731 | 0.00740*2 = 0.0148 |
| Interim 4 | 4 | 1100 | 1100/1750 = 0.629 | 0.01688*2 = 0.03376 | 2.19080 | 0.01423*2 = 0.0246 |
| Final | 5 | 1750 | 1750/1750 = 1 | 0.02500*2 = 0.05 | 2.19851 | 0.01396*2 = 0.02792 |

d) The authors present the results of the second interim analysis comparing SVR at 6 months on 700 patients. In the control group, 122 of the 350 patients had an SVR. In the new treatment group, 158 of the 350 patients had an SVR. Use a chi-square test to

determine the p-value for the test of:

$$H_0: p_{new\_treatment} = p_{control} \text{ vs. } H_A: p_{new\_treatment} \neq p_{control}$$

Based on the significance levels you calculated in Part D, should the DSMB recommend stopping the study for overwhelming efficacy of the new treatment? Explain briefly

We fail to reject the null at a p-value of 0.0055 with a significance level of 0.00548 and conclude that there is a not significant difference in the proportion of SVR events occurring in the treatment and control group. Hence, we should consider not stopping the study at the second interim analysis for overwhelming efficacy of the new treatment.The DSMB should not recommend stopping the trial