A randomized trial for Hepatitis C is planned with 1:1 randomization to a new treatment versus a standard of care treatment (control) is planned. The initial goal is to enroll 1500 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 1500 patients. Suppose that the first interim analysis is scheduled after the first 600 patients (300 per group) are treated and followed for 6 months, the second interim analysis is scheduled after the first 750 patients (375 per group) are treated and followed for 6 months, the third interim analysis is scheduled after the first 1000 patients (500 per group) are treated and followed for 6 months, the fourth interim analysis is scheduled after the first 1200 (600 per group) are treated and followed for 6 months, and the final analysis is planned after all 1500 patients have been treated and followed.

Question 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 the general difference between a significance level (at a given interim analysis) versus "alpha-spent" (at a given interim analysis).

O'Brien-Fleming Two-sided Significance Levels for the Five Analysis

n	s	$\alpha(s)$	Z	Significance Level
600	0.4	0.00079	3.3569	0.000788
750	0.5	0.00306	2.9885	0.00280
1,000	0.6667	0.01210	2.5396	0.01110
1,200	0.8	0.02442	2.3154	0.02058
1,500	1	0.05	2.0337	0.04198

Alpha-spent is the amount of the overall alpha, usually $\alpha = 0.05$, that is willing to be "spent" at time k. By the end of the experiment, the whole $\alpha = 0.05$ is willing to be spent. The significance level is the p-value associated with a critical chi-squared value that the experimental p-value is compared to determine if there is a significant difference between treatments.

Question B

It is now time for the first interim analysis comparing treatments. It turns out there are 560 total patients in this first analysis (280 per group) instead of the planned 600. 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 1500 patients and still planning for the remaining interim analyses to be carried out at the planned sample sizes discussed above (750, 1000, and 1200). In a nicely formatted table, please provide the alpha spent, critical values, significance levels at each interim stage and the final stage.

n	s	$\alpha(s)$	Z	Significance Level
560	0.3733	0.00048	3.4873	0.000488
750	0.5	0.00306	2.9790	0.00290
1,000	0.6667	0.01210	2.5391	0.01112
1,200	0.8	0.02442	2.3153	0.02060
1,500	1	0.05	2.0337	0.04198

Question 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 1500 to 1800 (900 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 560 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 750, 1000, and 1200 total patients, respectively, but that the final analysis will, again, be conducted on the 1800 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 1800 should match the alpha-spent by the second interim analysis under the old total sample size of 1500 in Part B, and similarly for analyses 3-5).

n	s	$\alpha(s)$	Z	Significance Level
560	0.3111	0.00048	3.4917	0.00048
750	0.4167	0.00306	2.9775	0.00290
1,000	0.5556	0.01210	2.5392	0.01112
1,200	0.6667	0.02442	2.3154	0.02060
1,800	1	0.05	2.0766	0.03784

Question D

Suppose the study is re-designed as in Part C, and now it is the time for the second interim analysis (again, the first interim analysis is already completed with a total sample size of 560). There are 822 total patients in this second interim analysis (411 in per group) instead of the planned 750. Re-calculate the significance levels to be used for this second interim analysis and in analyses 3-5, assuming analyses 3-5 will still be performed on 1000, 1200 and 1800 patients, respectively, and that the alpha-spent at the interim analyses should match what we calculated in Parts B & C.

n	s	$\alpha(s)$	Z	Significance Level
560	0.3111	0.00048	3.4917	0.00048
822	0.4567	0.00306	2.9832	0.00286
1,000	0.5556	0.01210	2.5290	0.01144
1,200	0.6667	0.02442	2.3133	0.02070
1,800	1	0.05	2.0762	0.03788

Question E

The authors present the results of the second interim analysis comparing SVR at 6 months on 822 patients. In the control group, 122 of the 411 patients had an SVR. In the new treatment group, 160 of the 411 patients had an SVR. Use a chi-square test to determine the p-value for the test of:

$$H_0: p_{new\ treatment} = p_{control}$$
$$H_A: p_{new\ treatment} \neq p_{control}$$

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

	No SVR	SVR	Total
Active Control	289 (70.32%)	122 (29.68%)	411
New Treatment	251 (61.07%)	160 (38.93%)	411
Total	540	282	822

A chi-squared test was used to test whether there was a difference in the proportion of sustained virological response between treatment groups. The chi-squared statistic was 7.7946 with 1 degree of freedom and the resulting p-value was 0.0052. With a p-value greater than the α=0.00286 O'Brien-Fleming significance level, the null hypothesis of there being no difference in in the proportion of sustained virological response between treatment groups was not rejected. There is insufficient statistical evidence to suggest early efficacy of the new treatment. The Data Safety and Monitoring Board should not recommend stopping the study early.