Feven Asefaha (fta2112)
Pharmaceutical Statistics (P8144)
Prof. Naitee Ting
March 12th, 2019
Homework #2

INTRODUCTION:

Within Phase 2 of clinical trials, dose-finding is a complex process. Investigators seek to demonstrate the Minimum Effective Dose (MinED), and the Maximum Tolerated Dose (MTD), in order to identify the most efficacious range for patients with a particular ailment. To this end, dose-response trials are necessary. By using three ordered levels of a drug, this study will compare each Dose level to a Placebo. In order to successfully make these comparisons, the study utilizes a multiple comparison procedure, a necessary part of reducing Type I Error. Along with the dose-response analysis, the paper takes an exploratory look at whether there is a linear trend within the dose levels and the response. In addition, there is an examination of potential interaction effects with patient sex and treatment, using stratification.

STATISTICAL HYPOTHESES

Primary:

For dose levels Low, Medium, and High, the proportion of those who exhibit a response in each group are being compared to the proportion of responders in the Placebo group. For these categorical comparisons, the hypotheses are the same at each dose level (Table 1):

Table 1. *Hypotheses for our primary dose-response comparisons*

Hypothesis	Low Dose	Medium Dose	High Dose
H_0	$p_{Treatment\_Responders} = p_{Placebo\_Responders}$		
H_A	$p_{Treatment\_Responders} \neq p_{Placebo\_Responders}$		

Alternately, the hypotheses could be written with respect to the Treatment and Placebo *Nonresponders*; since our outcome is binary, these are essentially the same thing. Because the goal is understanding efficacy, they are framed with respect to the Responders. These are two-sided hypotheses, so we are interested in any significant differences between the Placebo and treatment level. In order to evaluate these hypotheses, the Row Mean Scores produced by each comparison, along with their corresponding p-value, will be evaluated. The generalized hypotheses in SAS when looking at the Row Mean Scores of a j x k table are as follows:

$$H_0: The\ Row\ Mean\ Scores\ are\ Equal \quad vs. \quad H_A: The\ Row\ Mean\ Scores\ Differ$$

The Row Mean Score is an ANOVA statistic calculated from the comparison of ordinal treatment rows in the table, and a significant p-value will confirm whether there is a difference between the treatments.

Secondary:

To objectively test if there is a linear trend in response frequencies of an ordinal predictor variable, an overall trend test, and the subsequent Nonzero Correlation statistic, can be used. Similarly, this statistic uses the correlation r in its formulation. A significant p-value will explain a linear trend, although one must examine the frequency table produced in order to confirm direction (positive or negative). The generalized hypotheses in SAS when looking at the trend of a $j \times k$ table are:

H_0: *There is zero correlation*   vs. H_A: *There is a nonzero correlation*

Finally, to explore whether there is an interaction between treatment and sex, the Breslow-Day (B-D) Test for the Homogeneity of Odds Ratios will be used. The Breslow-Day Test uses a χ^2 distribution, and for each B-D Test in the study, the H_0 is that the Odds Ratios (ORs) between both sex and dose level are homogenous, or equal to each other. The stated hypotheses for the analysis are below (Table 2):

Table 2. *Hypotheses for the treatment and sex interaction analyses*

Hypothesis	Low Dose	Medium Dose	High Dose
H_0	$OR_{Female} = OR_{Male}$	$OR_{Female} = OR_{Male}$	$OR_{Female} = OR_{Male}$
H_A	$OR_{Female} \neq OR_{Male}$	$OR_{Female} \neq OR_{Male}$	$OR_{Female} \neq OR_{Male}$

For the B-D test, hypothesis test results are included in the output, but not the ORs per sex at each level. For transparency, ORs per sex were calculated for this study, using the stratified tables (Appendix, Figures I, K, M) and this formula:

$$Odds\ Ratio = \frac{\left(\frac{a}{c}\right)}{\left(\frac{b}{d}\right)} = \left(\frac{ad}{bc}\right) = \left(\frac{(PlaceboNonrespondersN) * (TreatmentRespondersN)}{(PlaceboRespondersN) * (TreatmentNonrespondersN)}\right)$$

Because the trend test and the sex & treatment interactions are considered exploratory in this study, no multiple comparison adjustments are used. Instead, the results may help inform future investigations of the study drug.

METHODS:

This is a multi-center study, with three dose levels, along with a Placebo group. Along with their ID and center, patient age and sex was also recorded for analysis. Each patient was assigned a dose level (Placebo, Low, Medium, or High). The outcome in this study is binary; patients are either categorized as a "Responder" or a "Non-Responder" to the dose they are given. All variables are described in Table 3.

Table 3. *Variable in the data, along with explanations*

Obs	Observation. Integer variable.
CTR	Center ID. Integer variable.
PID	Patient ID. Integer variable.
TRT	Describes treatment dose. Categorical variable. Coded as: 0 = Placebo, 1 = Low, 2 = Medium, 3 = High.
SEX	Describes patient's sex. Binary variable. Coded as: F = Female, M = Male
AGE	Patient's age, in years. Continuous variable.
RESP	Patient's response to treatment. Binary variable. Coded as: 0 = NON-RESPONDER , 1 = RESPONDER.

To compare the proportions from our dose groups to a placebo group, we have to control the Type I Error. Type I Error, also known as *alpha* (α), describes the percentage of false positives we are willing to accept. A false positive in this study would be seeing a difference between a treatment group and the Placebo group, when there is none. For one comparison in this study, $\alpha = 0.05$, so to compare just one of the dose groups to our placebo group, a 5% Type I error rate would be sufficient. But since there are three pairwise comparisons, we have to minimize our *alpha* threshold, so as to not continuously multiply, thus inflating, the chances of seeing group differences when there are none. Modifications to *alpha* must be made; these are known as Multiple Comparison Procedures (MCP). These procedures will control Type I Error, dependent on how many comparisons are needed. The MCP must be specified before the data is seen. For this study, the MCP will be Product Inequality. The modification is as follows:

$$1 - (1 - \alpha)^{\frac{1}{k}}, \text{ where } k \text{ refers to the number of comparisons.}$$
$$k = 3 \text{ (Placebo compared to Low, Medium, and High Doses)}$$
$$\alpha^* = 1 - \sqrt[3]{0.95} = \mathbf{0.0169}.$$

The response to a dose will be considered significantly different to the placebo response if the p-value is less than 0.0169. However, we still maintain $\alpha = 0.05$ for the linear trend and interaction tests, because these are not the primary aim of our study.

Pairwise comparisons, trend test, and interaction explorations were performed in SAS 9.4. Relevant SAS code and output are included in the Appendix. Base SAS code was provided by Professor Naitee Ting. For readability, SAS output was sometimes converted to tables using Microsoft Word 2016. Frequency tables were produced in Microsoft Excel 2016.

RESULTS:

This study enrolled 217 patients among 9 different centers, for a total of 121 women and 96 men. Each center enrolled between 16 to 35 patients. The average overall age among patients was 54.49 years. For each sex, average age was nearly equal (Appendix, Table 5). More breakdowns of patient demographics (including per center and per dose level), are included in the Appendix (Tables 4-11).

For the dose-response analyses, after the MCP adjustment, the Medium and High dose levels had a significantly different proportion of responders to treatment, as compared to Placebo. Our adjusted $\alpha = 0.0169$, so a significant proportion was not achieved at the Low dose level, compared to Placebo.

Table 4. Row Mean Scores, p-value, and ORs for dose-response analyses

Dose	Row Mean Scores	P-value	Odds Ratio
LOW	4.8915	0.0270	2.6033
MEDIUM	7.7239	0.0054	3.2727
HIGH	12.9321	0.0003	4.5818

Table 5 summarizes the results of the B-D interaction tests between male and female patients in the study.

Table 5. *Results for the three Breslow-Day Tests of Homogeneity between sexes, at dose levels Low, Medium, and High, as compared to Placebo*

Breslow-Day Test of Homogeneity of Odds Ratios						
Dose	$\widehat{OR}_{Female}$	$\widehat{OR}_{Male}$	H_0	H_A	χ^2	$Pr > \chi^2$
Low	2.347	3.378			0.1520	**0.6966**
Medium	3.059	3.897	$OR_{Female} = OR_{Male}$	$OR_{Female} \neq OR_{Male}$	0.0673	**0.7952**
High	4.225	6.333			0.1919	**0.6614**

No significant p-values were observed in any of the Breslow-Day Tests, indicating no difference in ORs between men and women at each dose level.

For the trend test, a significant p-value of 0.0004 demonstrated a correlated dose-response relationship within the data (Appendix, Figure H). Upon further examination of the output (Figure G), it is concluded that the correlation is positive, and the proportion of responders increases as the dose level increases. This is also supported by the increasing Row Mean Scores and decreasing p-values (Figures B, D, F).

DISCUSSION:

The Low Dose-Placebo comparison had a p-value of 0.0270, which was above our adjusted α threshold. Unlike other MCPs, Product Inequality does not assume a monotonic relationship between doses and responses. While we did discover monotonicity in the trend test, it was not required for our adjustment.

Because there were 1.26 times as many women as men, there were some dose levels at some centers that did not have any women at all. For example, Centers 506 and 508 did not have any men at the Medium dose level, and Center 507 did not have any male patients at the Low Dose level (Appendix, Table 11). While these analyses do not also stratify by center, a future attempt may be complicated by missing participants.

It is interesting that that in at all three Dose levels, no significant difference was found between the sexes. Men at all three treatment levels have greater odds of being a responder compared to placebo than women, most notably at the High Dose level. By this observation, men would appear to have a bigger effect size in Responders, compared to Non-Responders at the Low dose level. But recall that in our primary comparison, there was not a significant overall difference in the proportion of responders in the low dose group compared to the proportion of responders in the placebo group, which may explain why this sex difference is also not significant at this level.

CONCLUSION:

For our primary outcome, the effect size increased concurrently with the dose. After applying our MCP, Product Inequality, the Medium dose and High Dose groups both have a significantly different proportion of Responders, as compared to the Placebo group. For the linear trend test, a positive linear trend was observed, meaning the effect size increased as the dose increased. No interaction between treatment and sex was observed, as the Odds Ratios between sexes in each dose level were considered homogenous.

APPENDIX

Table 4. *Patients by Sex*

Sex	Total Patients
F	121
M	96
Grand Total	217

Table 5. *Average Age of Patients, by Sex*

SEX	Average Age
Female	54.45
Male	54.54
Overall	54.49

Table 6. *Total Patients per Center*

Center	Patients per Center
501	21
502	35
503	22
504	34
505	23
506	16
507	22
508	26
509	18
Grand Total	217

Table 7. *Average Patient Age, per Center*

Center	Average Age
501	55.71
502	52.94
503	56.05
504	54.47
505	53.57
506	55.50
507	56.27
508	53.19
509	54.17
Overall	54.49

Table 8. *Total Patients per Sex per Center*

Center	Sex	Patients
501	F	8
	M	13
	Total	21
502	F	18
	M	17
	Total	35
503	F	12
	M	10
	Total	22
504	F	17
	M	17
	Total	34
505	F	13
	M	10
	Total	23
506	F	9
	M	7
	Total	16
507	F	14
	M	8
	Total	22
508	F	19
	M	7
	Total	26
509	F	11
	M	7
	Total	18
	GRAND TOTAL	217

Table 9. *Average Age per Sex per Center*

CENTER	FEMALE	MALE
501	57.38	54.69
502	51.72	54.24
503	57.92	53.80
504	54.24	54.71
505	54.15	52.80
506	59.56	50.29
507	56.64	55.63
508	50.58	60.29
509	53.36	55.43

Table 10. *Patient total per Dose level per Center*

	Dose Levels				
Center	Placebo	Low	Medium	High	Total
501	5	5	6	5	21
502	9	9	8	9	35
503	6	5	5	6	22
504	9	8	9	8	34
505	6	6	5	6	23
506	4	4	4	4	16
507	5	6	6	5	22
508	7	7	6	6	26
509	5	4	5	4	18
Grand Total	56	54	54	53	217

Table 11. *Average Patient Age per Dose, per Sex, per Center. For the empty yellow cells, no male patients were enrolled at that dose level for that center.*

Center	Sex	DOSE			
		Placebo	Low	Medium	High
501	F	46.00	64.00	62.00	53.00
	M	56.67	55.33	56.00	51.75
502	F	50.63	55.00	50.75	51.50
	M	40.00	52.40	57.75	55.57
503	F	61.00	61.33	52.67	54.50
	M	54.00	57.50	54.50	51.50
504	F	59.83	49.00	52.67	53.33
	M	55.33	53.67	53.50	56.40
505	F	56.50	55.25	52.33	53.25
	M	54.50	58.50	50.00	46.50
506	F	74.00	61.00	55.75	59.33
	M	45.33	53.00		57.00
507	F	63.33	57.33	59.67	40.00
	M	59.50		50.33	58.33
508	F	59.25	46.40	52.67	44.00
	M	55.67	65.50		62.00
509	F	50.50	53.00	53.25	59.50
	M	52.00	50.67	54.00	65.00

MULTIPLE COMPARISON RESULTS

Figure A. *2 x 2 table, Response by Treatment (Low vs. Placebo)*

The FREQ Procedure

Frequency Row Pct	Table of TRT by RESP			
		RESP		
	TRT	0	1	Total
	0	45 80.36	11 19.64	56
	1	33 61.11	21 38.89	54
	Total	78	32	110

Figure B. *Statistic 2 and the accompanying p-value allows us to reject the H_0 that the Row Mean Scores between the proportion of Responders in the Low Dose and Placebo groups are the same.*

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	4.8915	0.0270
2	Row Mean Scores Differ	1	4.8915	0.0270
3	General Association	1	4.8915	0.0270

Figure C. *2 x 2 table, Response by Treatment (Medium vs. Placebo)*

The FREQ Procedure

Frequency Row Pct	Table of TRT by RESP			
		RESP		
	TRT	0	1	Total
	0	45 80.36	11 19.64	56
	2	30 55.56	24 44.44	54
	Total	75	35	110

Figure D. *Statistic 2 and the accompanying p-value allows us to reject the H_0 that the Row Mean Scores between the proportion of Responders in the Medium Dose and Placebo groups are the same.*

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	7.7239	0.0054
2	Row Mean Scores Differ	1	7.7239	0.0054
3	General Association	1	7.7239	0.0054

Figure E. *2 x 2 table, Response by Treatment (High vs. Placebo)*

The FREQ Procedure

Frequency Row Pct	Table of TRT by RESP			
		RESP		
	TRT	0	1	Total
	0	45 80.36	11 19.64	56
	3	25 47.17	28 52.83	53
	Total	70	39	109

Figure F. *Statistic 2 and the accompanying p-value allows us to reject the H_0 that the Row Mean Scores between the proportion of Responders in the High Dose and Placebo groups are the same.*

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	12.9321	0.0003
2	Row Mean Scores Differ	1	12.9321	0.0003
3	General Association	1	12.9321	0.0003

TREND TEST RESULTS

Figure G. *Trend test table. The frequency and percentage of responders increases with the dose.*

The FREQ Procedure

Frequency Row Pct	Table of TRT by RESP		
		RESP	
TRT	0	1	Total
0	45 80.36	11 19.64	56
1	33 61.11	21 38.89	54
2	30 55.56	24 44.44	54
3	25 47.17	28 52.83	53
Total	133	84	217

Figure H. *Statistic 1 and the associated p-value allow us to reject the H_0 and conclude that there is a nonzero correlation between dose and response.*

Statistic	Alternative Hypothesis	DF	Value	Prob
1	Nonzero Correlation	1	12.7349	0.0004
2	Row Mean Scores Differ	3	13.7208	0.0033
3	General Association	3	13.7208	0.0033

BRESLOW-DAY TEST RESULTS

Figure I. *2 x 2 tables, Response by Treatment (Low vs. Placebo), stratified by Sex*

The FREQ Procedure

Frequency Row Pct	Table 1 of TRT by RESP		
	Controlling for SEX=F		
		RESP	
TRT	0	1	Total
0	26 76.47	8 23.53	34
1	18 58.06	13 41.94	31
Total	44	21	65

Frequency Row Pct	Table 2 of TRT by RESP		
	Controlling for SEX=M		
		RESP	
TRT	0	1	Total
0	19 86.36	3 13.64	22
1	15 65.22	8 34.78	23
Total	34	11	45

Figure J. *The p-value does not let us reject H_0, and so there is homogeneity of the ORs between men and woman at the Low Dose level.*

Breslow-Day Test for Homogeneity of the Odds Ratios	
Chi-Square	0.1520
DF	1
Pr > ChiSq	0.6966

Figure K. *2 x 2 table, Response by Treatment (Medium vs. Placebo), stratified by Sex.*

The FREQ Procedure

Frequency Row Pct	Table 1 of TRT by RESP		
	Controlling for SEX=F		

TRT	RESP 0	1	Total
0	26 76.47	8 23.53	34
2	17 51.52	16 48.48	33
Total	43	24	67

Frequency Row Pct	Table 2 of TRT by RESP		
	Controlling for SEX=M		

TRT	RESP 0	1	Total
0	19 86.36	3 13.64	22
2	13 61.90	8 38.10	21
Total	32	11	43

Figure L. *The p-value does not let us reject H_0, and so there is homogeneity of the ORs between men and woman at the Medium Dose level.*

Breslow-Day Test for Homogeneity of the Odds Ratios	
Chi-Square	0.0673
DF	1
Pr > ChiSq	0.7952

Figure M. *2 x 2 table, Response by Treatment (High vs. Placebo), stratified by Sex*

The FREQ Procedure

Frequency Row Pct	Table 1 of TRT by RESP		
	Controlling for SEX=F		

TRT	RESP 0	1	Total
0	26 76.47	8 23.53	34
3	10 43.48	13 56.52	23
Total	36	21	57

Frequency Row Pct	Table 2 of TRT by RESP		
	Controlling for SEX=M		

TRT	RESP 0	1	Total
0	19 86.36	3 13.64	22
3	15 50.00	15 50.00	30
Total	34	18	52

Figure N. *The p-value does not let us reject H_0, and so there is homogeneity of the ORs between men and woman at the High Dose level.*

Breslow-Day Test for Homogeneity of the Odds Ratios	
Chi-Square	0.1919
DF	1
Pr > ChiSq	0.6614

SAS CODE

```sas
/* Pairwise test of low dose vs placebo */
proc freq data=dta41; tables trt*resp/cmh nopercent nocol; where trt in (0,1) ;

/* Pairwise test of medium dose vs placebo */
proc freq data=dta41; tables trt*resp/cmh nopercent nocol; where trt in (0,2) ;

/* Pairwise test of high dose vs placebo */
proc freq data=dta41; tables trt*resp/cmh nopercent nocol; where trt in (0,3) ;




/* Overall trend test (using /cmh)*/

proc freq data=dta41; tables trt*resp/cmh nopercent nocol ;




/* Treatment by sex interaction, using pairwise test of low dose vs placebo */
proc freq data=dta41; tables sex*trt*resp/cmh nopercent nocol; where trt in (0,1) ;

/* Treatment by sex interaction, using pairwise test of medium dose vs placebo */
proc freq data=dta41; tables sex*trt*resp/cmh nopercent nocol; where trt in (0,2) ;

/* Treatment by sex interaction, using pairwise test of high dose vs placebo */
proc freq data=dta41; tables sex*trt*resp/cmh nopercent nocol; where trt in (0,3) ;
```