Klyne Ludewick 19037000
Polawat Srichana 20696984

Question 1

The implementation of a factorial design in antiviral studies is introduced by the work done by Ding and his peers. The work was conducted by researchers looking to use a factorial design to see how five drugs affected the Herpes simplex virus type 1 (HSV -1). The purpose of the experiment conducted was to evaluate the antiviral effect of various drug combinations and to detect any interactions between the drugs to lower the percentage of viral infected cells. The different drugs are the experimental factors in this experiment. These are the three protein drugs Interferon-alpha (A), Interferon-Beta (B), Interferon-gamma (C), and the two chemical drugs Ribavirin (D), Acyclovir (E). The factorial design of this experiment contains a combination of two different designs. The researchers conducted two separate factorial designs a 16-run half factorial design with two levels and an 18-run orthogonal array with three levels. In the 16- run half factorial design each drug was tested at two levels a high (+1) or low (-1). The low value is defined as no drug used and the high value is a predetermined value of the drug in ng/mL. The high value in which the drug was used was determined by testing done prior to the experiment in which the drug is effective. In the 18-run orthogonal array, each drug was tested in three values a high, low, and mid-level. The mid-level in this experiment represented a diluted solution of the drugs used. The response in this experiment is the percentage of virus-infected cells after the treatment, this is referred to as the readout value. Having a low readout value is favourable as this represents how successful the drug is in attacking the virus. The treatment for this experiment tested all main effects of the drugs with no interaction, and then the following interactions AB, AC, AD, AE, BC, BD, BE CD, CE, and DE. These were the various combinations that were tested during the experiment and varying levels of high, low and mid treatments depending on the type of experiment performed. The principles of design were upheld during the design of this experiment. Randomization was implemented as the well plates were randomly assigned for a drug to be added to the plate. In addition, the experiment was conducted in a random order removing any factors an order could have on the number of virus-infected cells. Replication was implemented with two replications were done for each run. Blocking was achieved with the two researchers carrying out the experiment independently. The nuisance factors this blocked out is the experience gap between the two researchers, any pipetting error caused by researchers, and any mixing of the drug errors.

(Word Count 435)

Question2

We first examined the data obtained from Table 1. The results showed that in Table A
It consists of 5 drug variants: A = Interferon-alpha, B = Interferon-beta, C = Interferon-gamma, D = Ribavirin, E = Acyclovir, where A, B, C belong to protein-based drugs and D, E is the chemical drugs. There is also a level in which different drugs are divided into 3 levels (low, mid, high), as shown in the table. The design used contained a composite design of a two-level factorial design and a three-level orthogonal array. A half factorial design was

implemented as the number of runs required to perform a full 2^5 factorial would be costly. A second designed implement is a three-level orthogonal array, a feature of orthogonal arrays is that they are good at predicting parameters. This allows for a first and second-order model created based on the data collected during the experiment. In conclusion, a 34-run design is constructed by joining the factorial and orthogonal design. The design begins with creating a data frame like a Table B and then ordering replicate1, replicate2, respectively with a high or low level. After further investigation into the data, it was concluded a second order model was implemented with is squared variables. These squared variables were then imputed into the data frame. Finally, anova was conducted along with residuals vs fitted, and normal Q-Q to check that the assumptions for anova are upheld. These assumptions are normality, constant variance, and independence based of the graphs plotted it can be determined that these assumptions are upheld. Next, use a linear regression model was fitted to the data to achieve the same results conducted in the experiment by Ding et.al (2013) (Table C & Table D). Subsequently, to generate results based on column(b) according to model 3.1 fitting, the data transformation is required (Ding et. Al, 2013). We refit the data with a second-order model $y = \sqrt{readout}$ and the estimates are given in Table F & E. Finally, to fit the data according to the work done by Ding et. Al (2013) in Table H another linear model is required while removing replicate 1 and 14. The linear model fitted along with the parameters performed during our analysis are seen in Table G & H.

Interpretation

The contour graphs produced during the experiment represents the relationship between each drugs used. During each relationship plotted all values of the alternative drugs are held at a middle level (0). From graphs A and B, it can be interpreted that when A, B and C have drug dosage at high levels (1,1). It makes the virus infection percentage readings inappropriate. It has a very high percentage of viral infection and very high drug use (At all high levels the result of A, B and A, C are 22 and 25, respectively. Therefore, the minimum viral infection can be achieved by reducing A to low level (-1) while B and C is set as a high level (1). It means that overdose of A would interfere with the B and C efficiency of antiviral treatment. Graph C depicts the viral infection percentage when D and E are both at high drug levels. The readout result is approximately 5 so that's a minimum viral infection but if we reduce both D and E so we can indeed achieve the same antiviral readout. For example, reduce D and E to (0.5,0.5) the readout is approximately 10, moreover, the readout is lower than the interaction of A, B, and C at the same level. Lowering drug dosage can lead to reduced toxicity. Therefore, inhibiting the same amount of viral infection using a lower dosage of both D and E is practically meaningful.

This study showed that interactions within the interferon group (A, B, C) and chemical drug group (D, E) were significant. But the interaction between the two groups was insignificant. The interaction and quadratic effect of ribavirin and acyclovir showed that the dose had no effect on the antiviral effect. The relationship between the dose and the antiviral effect, and the trial was able to determine the optimal antiviral efficacy as well. Although lower dosages are used for both Ribavirin and Acyclovir when used as a combination. The application of an experimental design to investigate complex drug-drug interactions represents a new approach in virology. The limitation with this study is that not all interactions between drugs combinations were investigated. This was due to the limited number of resources available, and the time required to perform the experiment. As the

experiment look to investigate the interactions of drugs between had on lowering the virus. Further investigation into more combinations of drugs should be required to study the effects. Having more possible combinations of drugs used will help in lowering toxicity and resistant mutation in the virus.

In this study, we used 34 composite designs to study five drugs at three dose levels. According to the number of pills If the drug level is increased Larger and more complex designs were needed. This study, it shows that experimental design and analysis play an important role in studying more complex biological systems. From the, an increased number of drugs (more than 5 drugs) and a greater variety of drug levels (more than 3 levels) increases the composite design (more than 34-run) (Ding et. al 2013). This experiment had to be redesigned in a different way (alternative method). The reason is different variables will not produce results that are as effective as this experiment. More sophisticated experiments are needed to comply with complicated biological systems.

(Word Count 928)

Question 3

An alternative experimental design that can be used to model for this experiment is a central composite design. A central composite design will allow for a response surface model to be fitted from the data gathered. When determining an alternative design, it is important to remember the task of the experiment which is to study the effect that different drug combinations have on the readout value of the viral infection as the response variable. Using a 2^{5-1} factorial design we get 16 runs with 10-star points and 5 runs central point's making the experiment 32 runs in total. We have opted for a central composite design as we are able to fit a quadratic equation to find the optimal combinations of drug and their interactions to suggest a treatment. Having different combinations of drugs and finding the combinations that are statistically significant along with their interactions will aid doctors in prescribing medication to patients lowering the resistance mutation of the virus and the toxicity of various drugs. This design will be coded Low value (-1) or no drug, mid (0) and, a high (+1) maintaining the drug levels recommended by Ding et al. (2013). We have opted to run 5 trails at the centre point as this will help decrease the amount of variance within our model. Fitting this model will allow application into the optimum combination of drugs in which the percentage of viral infection is minimized. This design considers the limitation of the previous design in that there is a limited amount of money and available time to culture bacteria used in the previous experiment. Maintaining the design of experiment criteria this alternative design will do 2 replications, randomize the plates assigned to drug dosages, and will block by the researcher. Blocking by the researcher will limit the variation in each readout value due to pipetting and mixing errors that can be caused by the researcher. In conclusion, a response surface model will be a model in which a second-order model will be fitted to show the various interaction and find the optimum values of each drug combination in which readout is to minimize.

(Word Count 356)

Appendix

Table I. Factors and levels of the antiviral drug experiment

	Levels		
Factor	Low (−1)	Mid (0)	High (+1)
A = Interferon-alpha	no drug	1.56 ng/mL	50 ng/mL
B = Interferon-beta	no drug	1.56 ng/mL	50 ng/mL
C = Interferon-gamma	no drug	1.56 ng/mL	50 ng/mL
D = Ribavirin	no drug	781 ng/mL	25,000 ng/mL
E = Acyclovir	no drug	156 ng/mL	5000 ng/mL

Table A

B <dbl>	B2 <dbl>	C <dbl>	C2 <dbl>	D <dbl>	D2 <dbl>	E <dbl>	E2 <dbl>	Replicate1 <dbl>	Replicate2 <dbl>
-1	1	-1	1	-1	1	-1	1	69.8	72.0
1	1	-1	1	-1	1	-1	1	66.4	67.4
-1	1	1	1	-1	1	-1	1	83.0	68.6
-1	1	-1	1	1	1	-1	1	16.2	23.4
-1	1	-1	1	-1	1	1	1	46.1	33.6
1	1	1	1	-1	1	-1	1	68.6	65.5
1	1	-1	1	1	1	-1	1	6.8	7.2
1	1	-1	1	-1	1	1	1	15.6	19.1
-1	1	1	1	1	1	-1	1	11.1	7.0
-1	1	1	1	-1	1	1	1	19.8	20.3

1-10 of 34 rows | 4-13 of 13 columns

Previous 1 2 3 4 Next

Table B

```
Call:
lm(formula = a ~ A + B + C + D + E + A2 + B2 + C2 + D2 + E2 +
    A:B + A:C + A:D + A:E + B:C + B:D + B:E + C:D + C:E + D:E,
    data = etch_data)

Residuals:
    Min      1Q  Median      3Q     Max
-9.7956 -3.3077 -0.8241  2.5039 17.4987

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  17.3860     2.8455   6.110 1.83e-07 ***
A            -1.6042     0.8465  -1.895  0.06424 .
B            -2.3636     0.8465  -2.792  0.00755 **
C            -2.0136     0.8465  -2.379  0.02149 *
D           -19.5888     0.8465 -23.141  < 2e-16 ***
E           -13.9476     0.8465 -16.477  < 2e-16 ***
A2            3.6475     2.0678   1.764  0.08424 .
B2            0.1595     2.0678   0.077  0.93883
C2           -1.7042     2.0678  -0.824  0.41401
D2           -7.2142     2.0678  -3.489  0.00106 **
E2           15.3909     2.0678   7.443 1.75e-09 ***
A:B           0.7147     0.9259   0.772  0.44402
A:C           2.3356     0.9259   2.523  0.01509 *
A:D           1.4624     0.9259   1.580  0.12092
A:E          -1.3228     0.9259  -1.429  0.15970
B:C           1.6385     0.9259   1.770  0.08326 .
B:D          -0.2755     0.9259  -0.298  0.76732
B:E           1.3104     0.9259   1.415  0.16357
C:D          -0.6555     0.9259  -0.708  0.48246
C:E           0.1144     0.9259   0.124  0.90217
D:E           9.5616     0.9259  10.327 1.12e-13 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6.062 on 47 degrees of freedom
Multiple R-squared:  0.9602,    Adjusted R-squared:  0.9432
F-statistic: 56.62 on 20 and 47 DF,  p-value: < 2.2e-16
```

Table C

	(a)
Intercept	17.39***
A	−1.60
B	−2.36**
C	−2.01*
D	−19.59***
E	−13.95***
A^2	3.65
B^2	0.16
C^2	−1.70
D^2	−7.21**
E^2	15.39***
AB	0.71
AC	2.34*
AD	1.46
AE	−1.32
BC	1.64
BD	−0.28
BE	1.31
CD	−0.66
CE	0.11
DE	9.56***
Replicate	−0.52
$\hat{\sigma}$	6.095
R^2	0.961

Table D

```
Call:
lm(formula = EtchRate ~ A + B + C + D + E + A2 + B2 + C2 + D2 +
    E2 + A:B + A:C + A:D + A:E + B:C + B:D + B:E + C:D + C:E +
    D:E, data = etch_data1)

Residuals:
     Min       1Q   Median       3Q      Max
-0.99215 -0.32332 -0.04104  0.20805  1.58884

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.993474   0.256630  15.561  < 2e-16 ***
A           -0.125763   0.076345  -1.647  0.10617
B           -0.230798   0.076345  -3.023  0.00404 **
C           -0.201463   0.076345  -2.639  0.01125 *
D           -2.066413   0.076345 -27.067  < 2e-16 ***
E           -1.221764   0.076345 -16.003  < 2e-16 ***
A2           0.256812   0.186494   1.377  0.17502
B2           0.093066   0.186494   0.499  0.62009
C2          -0.006717   0.186494  -0.036  0.97142
D2          -1.171828   0.186494  -6.283 1.00e-07 ***
E2           1.412196   0.186494   7.572 1.11e-09 ***
A:B          0.124739   0.083502   1.494  0.14190
A:C          0.261193   0.083502   3.128  0.00302 **
A:D          0.075692   0.083502   0.906  0.36931
A:E         -0.125613   0.083502  -1.504  0.13919
B:C          0.142405   0.083502   1.705  0.09472 .
B:D         -0.085197   0.083502  -1.020  0.31281
B:E          0.129042   0.083502   1.545  0.12896
C:D         -0.107007   0.083502  -1.281  0.20631
C:E          0.050875   0.083502   0.609  0.54528
D:E          0.537736   0.083502   6.440 5.80e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5467 on 47 degrees of freedom
Multiple R-squared:  0.9646,    Adjusted R-squared:  0.9495
F-statistic: 64.03 on 20 and 47 DF,  p-value: < 2.2e-16
```

Table E

	(b)
Intercept	3.99***
A	−0.13
B	−0.23**
C	−0.20*
D	−2.07***
E	−1.22***
A^2	0.26
B^2	0.09
C^2	−0.01
D^2	−1.17***
E^2	1.41***
AB	0.12
AC	0.26**
AD	0.08
AE	−0.13
BC	0.14
BD	−0.09
BE	0.13
CD	−0.11
CE	0.05
DE	0.54***
Replicate	−0.03
$\hat{\sigma}$	0.5515
R^2	0.965

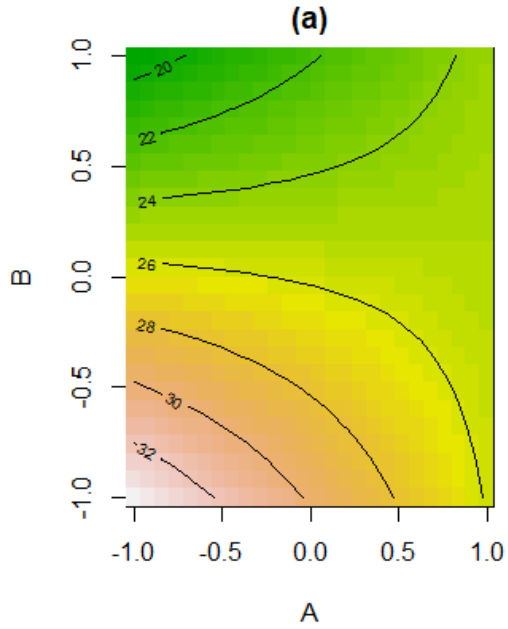
Table F

```
Call:
lm(formula = EtchRate ~ A + B + C + D + E + A2 + B2 + C2 + D2 +
    E2 + A:B + A:C + A:D + A:E + B:C + B:D + B:E + C:D + C:E +
    D:E, data = etch_data2)

Residuals:
     Min       1Q   Median       3Q      Max
-0.71756 -0.29014 -0.02907  0.20299  1.19336

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.075910   0.225375  18.085  < 2e-16 ***
A           -0.093250   0.067266  -1.386 0.172341
B           -0.271900   0.067569  -4.024 0.000211 ***
C           -0.224028   0.067002  -3.344 0.001652 **
D           -2.033900   0.067266 -30.237  < 2e-16 ***
E           -1.244330   0.067002 -18.571  < 2e-16 ***
A2           0.244109   0.163103   1.497 0.141313
B2           0.057433   0.163322   0.352 0.726703
C2          -0.046115   0.163378  -0.282 0.779011
D2          -1.184532   0.163103  -7.262 3.69e-09 ***
E2           1.372798   0.163378   8.403 7.59e-11 ***
A:B          0.174835   0.074117   2.359 0.022632 *
A:C          0.294491   0.073504   4.006 0.000223 ***
A:D          0.041619   0.073527   0.566 0.574122
A:E         -0.082451   0.073835  -1.117 0.269920
B:C          0.092692   0.074100   1.251 0.217298
B:D         -0.035101   0.074117  -0.474 0.638032
B:E          0.079328   0.074100   1.071 0.289956
C:D         -0.063845   0.073835  -0.865 0.391688
C:E          0.009834   0.073756   0.133 0.894515
D:E          0.571033   0.073504   7.769 6.50e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.478 on 46 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.9732,    Adjusted R-squared:  0.9616
F-statistic: 83.61 on 20 and 46 DF,  p-value: < 2.2e-16
```

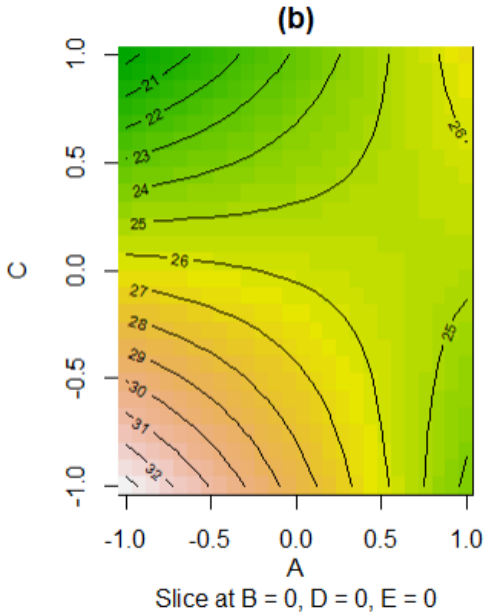
Table G

Intercept	(c)
A	4.08***
B	−0.09
C	−0.27***
D	−0.22**
E	−2.03***
A^2	−1.24***
B^2	0.24
C^2	0.06
D^2	−0.05
E^2	−1.18***
AB	1.37***
AC	0.17*
AD	0.29***
AE	0.04
BC	−0.08
BD	0.09
BE	−0.03
CD	0.08
CE	−0.06
DE	0.01
Replicate	0.57***
$\hat{\sigma}$	0.00
R^2	0.4833
	0.973

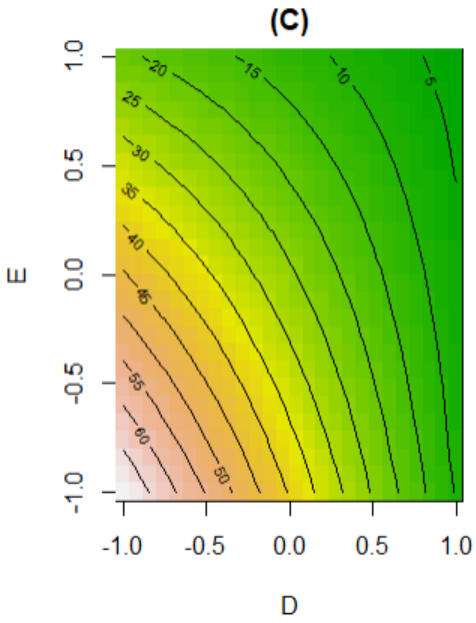
Table H



Graph A



Graph B



Graph C

References

Xianting Ding, Hongquan Xu,Chanelle Hopper,Jian Yangd and Chih-Ming Hoa (2013). Use of Fractional Factorial Designs in Antiviral Drug Studies.
https://onlinelibrary.wiley.com/doi/abs/10.1002/qre.1308