Modeling ECM1 Effect on Tumor Progression Using Generalized Linear Model

Yuanjie Jin
Department of Biology
Duke University
Durham, NC 27708
yj20@duke.edu

1 Introduction

Cancer is a type of malignant tumors whose initialization and progression are controlled by the malfunction of multiple genes. A major task in cancer research is to find out what genes are the major players in causing cancer, so that pharmaceutical companies and medical institutes can develop new medicines by targeting those genes to better treat cancer. Extracellular matrix protein 1 (ECM1), is one of such candidate genes. It was found from clinical data that the expression of ECM1 is frequently associated with cancer cells. For instance, ECM1 has a high expression level in a significant proportion of primary and secondary tumors including invasive breast ductal carcinoma (83%), esophageal squamous carcinoma (73%), gastric cancer (88%) and colorectal cancer (78%) [1]. Moreover, ECM1 is identified as a novel prognostic marker for poor long-term survival in breast carcinoma [2] and a predictor of the chemo-resistance in ovarian cancer patients [3]. ECM1 expression also correlates with the metastatic properties of tumors, with significant elevation in many malignant epithelial tumors that give rise to metastases [4]. These studies suggest that ECM1 likely play a role in tumor progression.

To examine the effect of ECM1 on tumor growth in vivo, McDonnell lab in Department of Pharmacology at Duke University recently conducted a study by knocking down ECM1 in mouse model. They used a particular line of cancer cells carrying shRNA (small hairpin RNA) that targets ECM1 for gene silencing and the shRNA can be induced by a drug doxycycline. These cells were injected into the breast of 8 week-old immune-compromised female mice on Day 0, such that the mice could develop breast cancer. Some of the mice (the test group) were fed with 2mg/ml doxycycline water starting three days prior to the cell injection to induce the expression of shECM1 and downregulate ECM1 level, whereas others (the control group) were fed with mock. The data from this experiment contains tumor volume and body weight measured every 2 or 3 days up until 5 weeks after injection, for a total number of 36 mice. It also includes relative ECM1 expression level within the tumor in each mouse measured at the end of the experiment, which verifies the knock-down of ECM1 in test group as supposed to the control group.

The plots of tumor volume vs. day from several representative mice indicate that there probably exists distinct tumor growth rates in different mice (Figure 1). The body weight, on the other hand, seems to remain largely stable during the period of time within each mice(Figure 2). It would be nice to have a quantitative measure of difference in tumor progression by some statistical methods and test whether ECM1 affects the difference in tumor growth and what is that effect. For this project, I was primarily interested in modeling ECM1 effect on the tumor growth rate using generalized linear models (GLMs). Since the experimental mice shared the same lineage and growth environment except that the ECM1 level is manipulated, I assume the tumor growth rates are a combination of a baseline value that is the same among mice and a variable value dependent on ECM1 level. Model comparison with analysis of deviance was performed to find the optimal model.

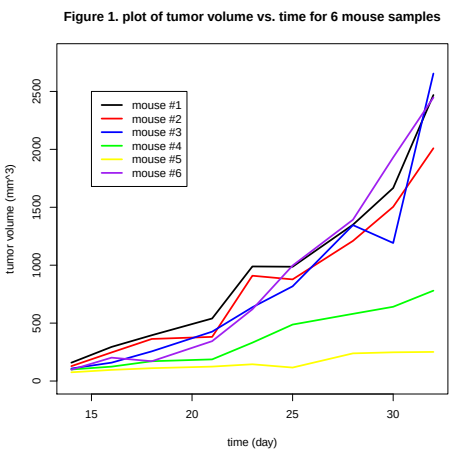

Figure 1: plot of tumor volume vs. time for 6 mouse samples.

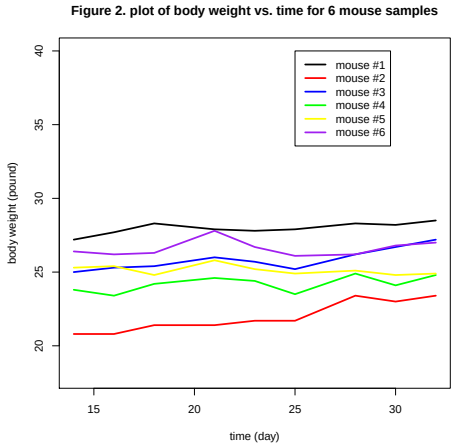

Figure 2: plot of body weight vs. time for 6 mouse samples.

2 Methods and Results

Let V_{it} be the tumor volume in ith mouse ($i = 1, \ldots, 36$) on Day t ($t \in \{14, 16, 18, 21, 23, 25, 28, 30, 32, 35\}$). As a starting point, V_{it} is modeled as independently Normally distributed:

$$V_{it} \sim \mathsf{N}(\mu_{it}, \sigma^2) \tag{1}$$

The plot of tumor volume vs. time(day) from several representative mice shows that tumor volume seems to increase exponentially by time in individual mice(Figure 1). Thus, an exponential growth model is applied to $E(V_{it})$:

$$\mu_{it} = E(V_{it}) = V_0 exp\{\alpha_i t\} \tag{2}$$

where V_0 is the tumor volume on Day 0 (because the same amount of cancer cells were injected into each mouse's breast on Day 0, I assume all mice have the same initial tumor volume), and α_i is the exponential growth constant of ith mouse. Let x_i be the ECM1 expression level in ith mouse, and W_{it} be the body weight of ith mouse on Day t. I first hypothesize that α_i has a linear relationship with x_i and potentially W_{i0}, the body weight measured on Day 0:

$$\alpha_i = \beta_0 + \beta_1 x_i + \beta_2 W_{i0} \tag{3}$$

Thus, a Gaussian family GLM (M0) is constructed with the link function

$$\log \mu_{it} = \log V_0 + \alpha_i t \tag{4}$$

$$= \log V_0 + \beta_0 t + \beta_1 x_i t + \beta_2 W_{i0} t \tag{5}$$

Shown below is a summary of model fitting results:

```
> gauss.m1 <- glm(formula = Tumor_volume ~ Day + ECM1xDay + Weight_0xDay,
family = gaussian(link = log), data = mouse.data)
> summary(gauss.m1)

Call:
glm(formula = Tumor_volume ~ Day + ECM1xDay + Weight_0xDay,
family = gaussian(link = log), data = mouse.data)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-1630.34   -125.23    -39.21    135.34   1197.41

Coefficients:
```

```
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.609e+00  1.502e-01  24.032  < 2e-16 ***
Day          7.922e-02  1.054e-02   7.519 4.52e-13 ***
ECM1xDay     2.779e-04  4.667e-05   5.955 6.23e-09 ***
Weight_0xDay 1.426e-03  4.108e-04   3.471 0.000583 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

(Dispersion parameter for gaussian family taken to be 132842.7)

    Null deviance: 187758430  on 359  degrees of freedom
Residual deviance:  47292105  on 356  degrees of freedom
AIC: 5274.5

Number of Fisher Scoring iterations: 6
```

From the residuals vs. fitted plot of this model (Figure 3), we can see the fitted values are clearly heteroscedastic (i.e., there is an increasing trend of residual deviance with increasing fitted values). One possibility for this heteroscedasticity could be that the tumor volume has a greater variation (σ^2) at a later stage, which is totally reasonable.

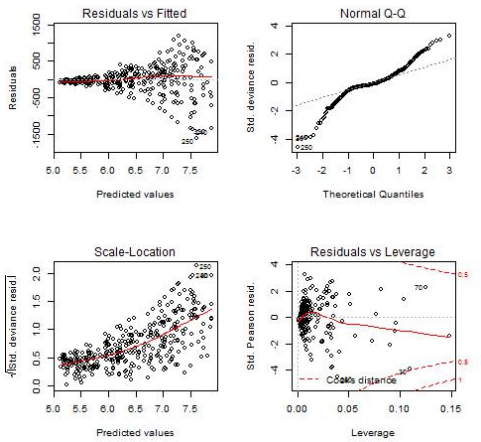

Figure 3: Diagnostic plots of Gaussian model M0 with an log link.

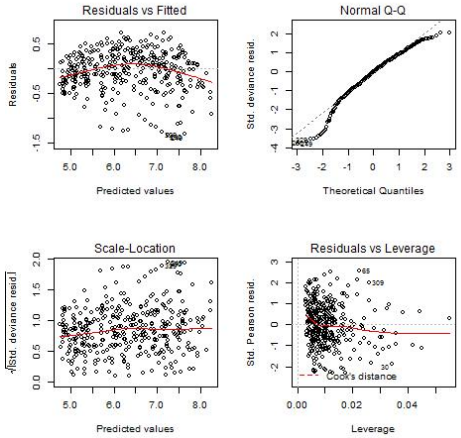

Figure 4: Diagnostic plots of Gamma model M3 with an log link.

To account for the predictor-dependent variance change, V_{it} is then modeled as a GLM with Gamma errors:

$$V_{it} \sim \mathsf{G}(k_{it}, \theta) \tag{6}$$

where k_{it} is the shape and θ is the scale. By the properties of Gamma distribution, we have

$$\mu_{it} = E(V_{it}) = k_{it}\theta \tag{7}$$

$$Var(V_{it}) = k_{it}\theta^2 = \mu_{it}\theta \tag{8}$$

from which we see that the variance of V_{it} is proportional to the mean of V_{it}.

This model (M3) is fitted with the same log link mentioned above, giving the following summary:

```
> gamma.m3 <- glm(formula = Tumor_volume ~ Day + ECM1xDay + Weight_0xDay,
family = Gamma(link = 'log'), data = mouse.data)
> summary(gamma.m3)

Call:
glm(formula = Tumor_volume ~ Day + ECM1xDay + Weight_0xDay,
```

```
family = Gamma(link = "log"), data = mouse.data)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-1.34035  -0.27261   0.00265   0.22436   0.72502

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.004e+00  7.045e-02  42.642  < 2e-16 ***
Day          9.294e-02  1.187e-02   7.831 5.61e-14 ***
ECM1xDay     5.601e-04  6.056e-05   9.248  < 2e-16 ***
Weight_0xDay 1.558e-03  5.116e-04   3.046  0.00249 **
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

(Dispersion parameter for Gamma family taken to be 0.1268675)

    Null deviance: 325.893  on 359  degrees of freedom
Residual deviance:  56.346  on 356  degrees of freedom
AIC: 4850.4

Number of Fisher Scoring iterations: 5

> confint(gamma.m3)
Waiting for profiling to be done...
                    2.5 %       97.5 %
(Intercept)  2.8619675536 3.1476623201
Day          0.0702174145 0.1158636356
ECM1xDay     0.0004352306 0.0006857878
Weight_0xDay 0.0005768177 0.0025344350
```

Note that the Gamma model largely addresses the issue of heteroscedasticity, as shown by its residuals vs. fitted plot (Figure 4). On the other hand, the assumption of linear relation between α_i and x_i or W_{i0} might be too simplistic. In particular, α_i may also involve an interaction term $x_i \times W_{i0}$. To explore such a possibility, I did a model comparison for the following models:

$$\text{M1}: \log \mu_{it} = \log V_0 + \beta_0 t \tag{9}$$

$$\text{M2}: \log \mu_{it} = \log V_0 + \beta_0 t + \beta_1 x_i t \tag{10}$$

$$\text{M3}: \log \mu_{it} = \log V_0 + \beta_0 t + \beta_1 x_i t + \beta_2 W_{i0} t \tag{11}$$

$$\text{M4}: \log \mu_{it} = \log V_0 + \beta_0 t + \beta_1 x_i t + \beta_2 W_{i0} t + \beta_3 x_i W_{i0} t \tag{12}$$

An analysis of deviance for these models is shown below:

```
> gamma.m1 <- glm(formula = Tumor_volume ~ Day, family = Gamma(link = 'log'), data =
> gamma.m2 <- glm(formula = Tumor_volume ~ Day + ECM1xDay, family = Gamma(link = 'loc
> gamma.m4 <- glm(formula = Tumor_volume ~ Day + ECM1xDay + Weight_0xDay + ECM1xWeigh
> anova(gamma.m1, gamma.m2, gamma.m3, gamma.m4, test = 'Chi')
Analysis of Deviance Table

Model 1: Tumor_volume ~ Day
Model 2: Tumor_volume ~ Day + ECM1xDay
Model 3: Tumor_volume ~ Day + ECM1xDay + Weight_0xDay
Model 4: Tumor_volume ~ Day + ECM1xDay + Weight_0xDay + ECM1xWeight_0xDay
  Resid. Df Resid. Dev Df Deviance  Pr(>Chi)
1       358     67.020
2       357     57.571  1   9.4494 < 2.2e-16 ***
3       356     56.346  1   1.2243  0.001753 **
4       355     55.931  1   0.4152  0.068411 .
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

from which we see

1) the linear terms of both x_i and W_{i0} are important in reducing the residual deviance in the models, with the former much more significant than the latter.

2) further inclusion of the interaction term of $x_i \times W_{i0}$ does not significantly (at a significance level of 0.05) explain much of the residual deviance.

Therefore, I decided to stick with the model M3 for parsimonious purpose. Then I tried to test the potential random effect on the intercept and β_0 conferred by grouping of mice in cages. The corresponding generalized mixed effect model (GLMM) was constructed as follows:

$$\text{M5} : \log \mu_{it} = \log V_0 + \tau_j + (\beta_0 + \gamma_j)t + \beta_1 x_i t + \beta_2 W_{i0} t \tag{13}$$

where $j = 1, \ldots, 8$ indexes the cage, $\tau_j \sim i.i.d. Normal(0, \sigma_\tau^2)$ is the random effect on the intercept and $\gamma_j \sim i.i.d. Normal(0, \sigma_\gamma^2)$ is the random effect on the coefficient of the covariate t. This GLMM failed to converge, however. Part of the reason could be the model M3 is already a little too complicated to include random effects. Moreover, the R program *glmer* (package 'lme4') warns that some predictor variables are on very different scales, so a rescaling of the data might help.

3 Conclusions and Implications

In this study, the Gamma family GLM with a log link fits the data better than the Gaussian model with the same link, because the Gamma model works well for positive-only data with positively-skewed errors. The findings from the modeling support the hypothesis that ECM1 promotes breast cancer progression in mice. Specifically, as the summary of model M3 shows, 1 unit increase in ECM1 level leads to $exp\{5.601 \times 10^{-4} t\} - 1$ fold increase in mean tumor volume on Day t. Meanwhile, the initial body weight on Day 0 also seems to have a positive effect on tumor growth, indicating that a healthier mouse is more likely to have a more aggressive tumor progression in this case.

References

[1] Wang, L., et al. (2003) Extracellular matrix protein 1 (ECM1) is over-expressed in malignant epithelial tumors. *Cancer Lett*, **200**(1): 57-67.

[2] Lal, G., et al. (2009) Extracellular matrix 1 (ECM1) expression is a novel prognostic marker for poor long-term survival in breast cancer: a Hospital-based Cohort Study in Iowa. *Ann Surg Oncol*, **16**(8): 2280-7.

[3] Pan, S., et al. (2009) Quantitative proteomics analysis integrated with microarray data reveals that extracellular matrix proteins, catenins, and p53 binding protein 1 are important for chemotherapy response in ovarian cancers. *OMICS* **13**(4):345-54.

[4] Chen, H., et al. (2011) Extracellular matrix protein 1, a novel prognostic factor, is associated with metastatic potential of hepatocellular carcinoma. *Medical Oncology* **28**:S318-S325.

Appendix

Raw data used in this project

```
Cage Mouse Day Weight Weight_0 Tumor_volume ECM1_level
1 1 14 27.2 24.8 158.37 27.34580427
1 1 16 27.7 24.8 294.58 27.34580427
1 1 18 28.3 24.8 395.45 27.34580427
1 1 21 27.9 24.8 539.97 27.34580427
1 1 23 27.8 24.8 989.04 27.34580427
1 1 25 27.9 24.8 986.7 27.34580427
1 1 28 28.3 24.8 1349.42 27.34580427
1 1 30 28.2 24.8 1666.21 27.34580427
1 1 32 28.5 24.8 2469.04 27.34580427
1 1 35 na 24.8 3141.67 27.34580427
1 2 14 23.7 21.6 171.11 16.31389181
1 2 16 23.4 21.6 192.94 16.31389181
```

```
1 2 18 23.9 21.6 313.82 16.31389181
1 2 21 23.4 21.6 597.69 16.31389181
1 2 23 23.7 21.6 601.64 16.31389181
1 2 25 23.6 21.6 691.17 16.31389181
1 2 28 24.3 21.6 1070.75 16.31389181
1 2 30 24.7 21.6 1424.66 16.31389181
1 2 32 24.7 21.6 1681.88 16.31389181
1 2 35 na 21.6 2556.52 16.31389181
1 3 14 24.9 22.2 127.05 40.84785209
1 3 16 24.9 22.2 173.71 40.84785209
1 3 18 25.4 22.2 188.48 40.84785209
1 3 21 24.6 22.2 350.89 40.84785209
1 3 23 24.2 22.2 405.22 40.84785209
1 3 25 24.1 22.2 448.08 40.84785209
1 3 28 24.7 22.2 571.28 40.84785209
1 3 30 25 22.2 772.25 40.84785209
1 3 32 24.7 22.2 897.25 40.84785209
1 3 35 na 22.2 1325.58 40.84785209
1 4 14 21.6 20.2 155.94 35.03055839
1 4 16 22 20.2 173.37 35.03055839
1 4 18 22.1 20.2 267.1 35.03055839
1 4 21 23 20.2 470.77 35.03055839
1 4 23 23.2 20.2 563.79 35.03055839
1 4 25 23.2 20.2 820.59 35.03055839
1 4 28 23.6 20.2 1103.46 35.03055839
1 4 30 23.6 20.2 1473.28 35.03055839
1 4 32 23.9 20.2 1896.79 35.03055839
1 4 35 na 20.2 2122.75 35.03055839
2 5 14 28 23.1 170.25 17.3826806
2 5 16 27.3 23.1 279.62 17.3826806
2 5 18 27.7 23.1 289.75 17.3826806
2 5 21 27.5 23.1 507.5 17.3826806
2 5 23 27.1 23.1 763.33 17.3826806
2 5 25 27.3 23.1 798.45 17.3826806
2 5 28 27.8 23.1 1319.59 17.3826806
2 5 30 27.6 23.1 1364.28 17.3826806
2 5 32 28.2 23.1 1769.24 17.3826806
2 5 35 na 23.1 2545.16 17.3826806
2 6 14 26.2 23.4 135.33 17.02043739
2 6 16 25.7 23.4 245.61 17.02043739
2 6 18 26.1 23.4 336.76 17.02043739
2 6 21 26.4 23.4 418.99 17.02043739
2 6 23 26.3 23.4 526.99 17.02043739
2 6 25 26.2 23.4 781.37 17.02043739
2 6 28 26.8 23.4 1089.76 17.02043739
2 6 30 26.8 23.4 1743.72 17.02043739
2 6 32 27 23.4 2094.99 17.02043739
2 6 35 na 23.4 2888.07 17.02043739
2 7 14 20.8 18.4 128.43 28.26223354
2 7 16 20.8 18.4 246.93 28.26223354
2 7 18 21.4 18.4 363.25 28.26223354
2 7 21 21.4 18.4 381.22 28.26223354
2 7 23 21.7 18.4 909.07 28.26223354
2 7 25 21.7 18.4 877.47 28.26223354
2 7 28 23.4 18.4 1209.78 28.26223354
2 7 30 23 18.4 1504.01 28.26223354
2 7 32 23.4 18.4 2009.38 28.26223354
2 7 35 na 18.4 2727.64 28.26223354
2 8 14 23.8 20.3 189.75 31.46316161
2 8 16 23.7 20.3 249.08 31.46316161
2 8 18 24.3 20.3 271.68 31.46316161
2 8 21 23.9 20.3 575.35 31.46316161
2 8 23 24.6 20.3 607.8 31.46316161
2 8 25 24.3 20.3 905.99 31.46316161
2 8 28 24.5 20.3 1264.75 31.46316161
```

```
2 8 30 24.5 20.3 1297.03 31.46316161
2 8 32 25.4 20.3 1894.06 31.46316161
2 8 35 na 20.3 2303.06 31.46316161
3 9 14 25.8 23.2 122.91 20.89459188
3 9 16 26.1 23.2 158.72 20.89459188
3 9 18 26.5 23.2 223.85 20.89459188
3 9 21 27.1 23.2 192.07 20.89459188
3 9 23 26.4 23.2 533.14 20.89459188
3 9 25 26 23.2 579.33 20.89459188
3 9 28 26.7 23.2 870.58 20.89459188
3 9 30 26.7 23.2 1123.89 20.89459188
3 9 32 26.8 23.2 1635.67 20.89459188
3 9 35 na 23.2 1868.86 20.89459188
3 10 14 26.6 23.4 129.71 20.46580111
3 10 16 25.8 23.4 245.74 20.46580111
3 10 18 26.2 23.4 358.98 20.46580111
3 10 21 27 23.4 500.64 20.46580111
3 10 23 26.1 23.4 901.13 20.46580111
3 10 25 25.9 23.4 1004.53 20.46580111
3 10 28 26.9 23.4 1537.04 20.46580111
3 10 30 27.2 23.4 1204.12 20.46580111
3 10 32 27.8 23.4 2677.07 20.46580111
3 10 35 na 23.4 3016.81 20.46580111
3 11 14 25.5 22.9 142.23 37.39140852
3 11 16 25.3 22.9 194.86 37.39140852
3 11 18 25.7 22.9 379.85 37.39140852
3 11 21 25.3 22.9 411.18 37.39140852
3 11 23 25.1 22.9 914.51 37.39140852
3 11 25 24.9 22.9 681.51 37.39140852
3 11 28 26.1 22.9 1305.12 37.39140852
3 11 30 25.4 22.9 1316.34 37.39140852
3 11 32 25.9 22.9 1662 37.39140852
3 11 35 na 22.9 1937.15 37.39140852
3 12 14 25.6 22.9 157.97 21.81694003
3 12 16 25.7 22.9 177.94 21.81694003
3 12 18 25.8 22.9 227.36 21.81694003
3 12 21 26.2 22.9 407.03 21.81694003
3 12 23 25.3 22.9 485.54 21.81694003
3 12 25 24.7 22.9 825.3 21.81694003
3 12 28 25.9 22.9 1299.12 21.81694003
3 12 30 26.2 22.9 1548.42 21.81694003
3 12 32 26.4 22.9 2266.89 21.81694003
3 12 35 na 22.9 2970.71 21.81694003
3 13 14 25 22.7 107.92 11.72030239
3 13 16 25.3 22.7 158.69 11.72030239
3 13 18 25.4 22.7 256.28 11.72030239
3 13 21 26 22.7 426.2 11.72030239
3 13 23 25.7 22.7 636.93 11.72030239
3 13 25 25.2 22.7 817.72 11.72030239
3 13 28 26.2 22.7 1345.99 11.72030239
3 13 30 26.7 22.7 1191.74 11.72030239
3 13 32 27.2 22.7 2654.58 11.72030239
3 13 35 na 22.7 2596.88 11.72030239
4 14 14 24.6 22.6 98.05 13.56428416
4 14 16 24.1 22.6 217.82 13.56428416
4 14 18 24.8 22.6 253.68 13.56428416
4 14 21 25.8 22.6 551.36 13.56428416
4 14 23 25.1 22.6 749.59 13.56428416
4 14 25 24.4 22.6 796.34 13.56428416
4 14 28 25.7 22.6 957.22 13.56428416
4 14 30 24.8 22.6 1274.67 13.56428416
4 14 32 25.5 22.6 2203.42 13.56428416
4 14 35 na 22.6 2302.02 13.56428416
4 15 14 25.3 23 146.91 28.270151
4 15 16 25.4 23 205.1 28.270151
```

4 15 18 25 23 426.3 28.270151
4 15 21 25.6 23 648.03 28.270151
4 15 23 25.1 23 810.15 28.270151
4 15 25 25.4 23 1038.28 28.270151
4 15 28 26.5 23 1242.76 28.270151
4 15 30 26 23 1572.2 28.270151
4 15 32 26.6 23 2573.71 28.270151
4 15 35 na 23 2354.28 28.270151
4 16 14 25.2 22.4 165.9 25.04223179
4 16 16 25.5 22.4 220.32 25.04223179
4 16 18 25.6 22.4 266.52 25.04223179
4 16 21 26.4 22.4 476.79 25.04223179
4 16 23 26.1 22.4 737.5 25.04223179
4 16 25 25.9 22.4 744.07 25.04223179
4 16 28 26.7 22.4 1148.86 25.04223179
4 16 30 26.2 22.4 1202.16 25.04223179
4 16 32 25.8 22.4 1457.97 25.04223179
4 16 35 na 22.4 2051.72 25.04223179
4 17 14 26.7 25.8 119.02 21.45774082
4 17 16 26.4 25.8 222.86 21.45774082
4 17 18 26.6 25.8 296.22 21.45774082
4 17 21 26.5 25.8 374.57 21.45774082
4 17 23 26.4 25.8 569.94 21.45774082
4 17 25 25.9 25.8 707.12 21.45774082
4 17 28 26.4 25.8 1089.37 21.45774082
4 17 30 25.8 25.8 975.16 21.45774082
4 17 32 26 25.8 1313.6 21.45774082
4 17 35 na 25.8 2153.87 21.45774082
4 18 14 25.2 22.3 137.62 39.04566824
4 18 16 25.1 22.3 267.89 39.04566824
4 18 18 25.4 22.3 369.57 39.04566824
4 18 21 26.2 22.3 381.97 39.04566824
4 18 23 25.9 22.3 834.75 39.04566824
4 18 25 25.4 22.3 617.41 39.04566824
4 18 28 26.7 22.3 1092.62 39.04566824
4 18 30 26.5 22.3 1476.39 39.04566824
4 18 32 26.7 22.3 1869.09 39.04566824
4 18 35 na 22.3 2259.52 39.04566824
5 19 14 23.8 22.8 99.75 5.08465358
5 19 16 23.4 22.8 124.26 5.08465358
5 19 18 24.2 22.8 171 5.08465358
5 19 21 24.6 22.8 186.87 5.08465358
5 19 23 24.4 22.8 331.24 5.08465358
5 19 25 23.5 22.8 487.45 5.08465358
5 19 28 24.9 22.8 580.13 5.08465358
5 19 30 24.1 22.8 641.08 5.08465358
5 19 32 24.8 22.8 780.25 5.08465358
5 19 35 na 22.8 1024.36 5.08465358
5 20 14 23.6 21.9 73.83 9.867789107
5 20 16 23.3 21.9 157.78 9.867789107
5 20 18 23.5 21.9 183.06 9.867789107
5 20 21 23.9 21.9 213.46 9.867789107
5 20 23 23.6 21.9 355.78 9.867789107
5 20 25 22.6 21.9 569.83 9.867789107
5 20 28 24 21.9 985.07 9.867789107
5 20 30 23.1 21.9 644.02 9.867789107
5 20 32 23.5 21.9 987.4 9.867789107
5 20 35 na 21.9 1282.63 9.867789107
5 21 14 25.2 21.7 103.29 12.59660406
5 21 16 25.1 21.7 174.29 12.59660406
5 21 18 25.3 21.7 159.77 12.59660406
5 21 21 25.9 21.7 154.21 12.59660406
5 21 23 25.6 21.7 356.56 12.59660406
5 21 25 24.7 21.7 510.56 12.59660406
5 21 28 25.8 21.7 926.29 12.59660406

5 21 30 25.2 21.7 667.39 12.59660406
5 21 32 25.7 21.7 1048.4 12.59660406
5 21 35 na 21.7 1305.78 12.59660406
5 22 14 23.9 21.3 97.91 7.095552573
5 22 16 24.1 21.3 109.19 7.095552573
5 22 18 24.2 21.3 197.02 7.095552573
5 22 21 25.4 21.3 171.73 7.095552573
5 22 23 24.8 21.3 359.11 7.095552573
5 22 25 24.8 21.3 614.94 7.095552573
5 22 28 25.3 21.3 892.33 7.095552573
5 22 30 24.9 21.3 1115.1 7.095552573
5 22 32 25.3 21.3 1404.71 7.095552573
5 22 35 na 21.3 1532.93 7.095552573
6 23 14 21.3 19.1 81.61 1.21708907
6 23 16 20.7 19.1 84.08 1.21708907
6 23 18 20.9 19.1 132.92 1.21708907
6 23 21 22.2 19.1 238.02 1.21708907
6 23 23 21.8 19.1 313.04 1.21708907
6 23 25 21.7 19.1 254.82 1.21708907
6 23 28 22.3 19.1 435.94 1.21708907
6 23 30 22.3 19.1 639.49 1.21708907
6 23 32 22.1 19.1 891.15 1.21708907
6 23 35 na 19.1 912.6 1.21708907
6 24 14 24.5 22.8 79.87 2.095994555
6 24 16 24.2 22.8 82.5 2.095994555
6 24 18 24 22.8 153.81 2.095994555
6 24 21 25.1 22.8 167.21 2.095994555
6 24 23 24.5 22.8 255.53 2.095994555
6 24 25 24.7 22.8 225.89 2.095994555
6 24 28 25.6 22.8 434.22 2.095994555
6 24 30 25.1 22.8 441.79 2.095994555
6 24 32 25.5 22.8 702.54 2.095994555
6 24 35 na 22.8 501.29 2.095994555
6 25 14 25.3 23.7 75.43 4.80631421
6 25 16 25.4 23.7 96.2 4.80631421
6 25 18 24.8 23.7 110.95 4.80631421
6 25 21 25.8 23.7 124.59 4.80631421
6 25 23 25.2 23.7 144.67 4.80631421
6 25 25 24.9 23.7 116.38 4.80631421
6 25 28 25.1 23.7 238.78 4.80631421
6 25 30 24.8 23.7 247.99 4.80631421
6 25 32 24.9 23.7 251.46 4.80631421
6 25 35 na 23.7 390.83 4.80631421
6 26 14 23.3 21.7 88.17 3.392334562
6 26 16 23 21.7 105.78 3.392334562
6 26 18 23.3 21.7 181.71 3.392334562
6 26 21 24 21.7 219.8 3.392334562
6 26 23 23.6 21.7 343.34 3.392334562
6 26 25 23.2 21.7 394.29 3.392334562
6 26 28 23.9 21.7 556.55 3.392334562
6 26 30 23.4 21.7 605.91 3.392334562
6 26 32 23.4 21.7 925.39 3.392334562
6 26 35 na 21.7 762.84 3.392334562
7 27 14 25.8 22.6 86.55 5.218424714
7 27 16 25.2 22.6 151.41 5.218424714
7 27 18 25.3 22.6 145 5.218424714
7 27 21 25.4 22.6 291.54 5.218424714
7 27 23 24.9 22.6 508.35 5.218424714
7 27 25 24.6 22.6 909.49 5.218424714
7 27 28 25.2 22.6 1194.3 5.218424714
7 27 30 24.6 22.6 1318.85 5.218424714
7 27 32 24.3 22.6 1801.77 5.218424714
7 27 35 na 22.6 2688.86 5.218424714
7 28 14 24.5 22.8 97.98 4.601780041
7 28 16 24 22.8 148.09 4.601780041

```
7 28 18 23.8 22.8 157.03 4.601780041
7 28 21 24.6 22.8 268.39 4.601780041
7 28 23 24 22.8 412.02 4.601780041
7 28 25 23.7 22.8 915.7 4.601780041
7 28 28 23.9 22.8 961.53 4.601780041
7 28 30 24.5 22.8 1270.08 4.601780041
7 28 32 25 22.8 1755.52 4.601780041
7 28 35 na 22.8 2775.67 4.601780041
7 29 14 25.1 22.8 142.51 1.779282023
7 29 16 25.7 22.8 169.15 1.779282023
7 29 18 25 22.8 199.64 1.779282023
7 29 21 26.1 22.8 358.82 1.779282023
7 29 23 25.7 22.8 503.69 1.779282023
7 29 25 25 22.8 606.52 1.779282023
7 29 28 25.5 22.8 923.56 1.779282023
7 29 30 25.5 22.8 1329.04 1.779282023
7 29 32 26.1 22.8 1620.66 1.779282023
7 29 35 na 22.8 2435.7 1.779282023
7 30 14 22.5 20.3 130.17 1.339453517
7 30 16 22.6 20.3 148.56 1.339453517
7 30 18 22.3 20.3 220.8 1.339453517
7 30 21 23.6 20.3 267.77 1.339453517
7 30 23 23 20.3 363.67 1.339453517
7 30 25 22.8 20.3 685.93 1.339453517
7 30 28 23.2 20.3 906.53 1.339453517
7 30 30 23.9 20.3 1278.53 1.339453517
7 30 32 23.7 20.3 1680.67 1.339453517
7 30 35 na 20.3 2013.35 1.339453517
7 31 14 26.4 25.1 97.75 2.197434664
7 31 16 26.2 25.1 201.75 2.197434664
7 31 18 26.3 25.1 170.84 2.197434664
7 31 21 27.8 25.1 344.24 2.197434664
7 31 23 26.7 25.1 619.29 2.197434664
7 31 25 26.1 25.1 997 2.197434664
7 31 28 26.2 25.1 1392.61 2.197434664
7 31 30 26.8 25.1 1930.25 2.197434664
7 31 32 27 25.1 2447.35 2.197434664
7 31 35 na 25.1 2870.51 2.197434664
8 32 14 24.7 22.7 116.58 1
8 32 16 24.4 22.7 132.78 1
8 32 18 24.8 22.7 208.45 1
8 32 21 25.5 22.7 264.92 1
8 32 23 25.2 22.7 404.36 1
8 32 25 24.6 22.7 485.86 1
8 32 28 25.7 22.7 1024.39 1
8 32 30 25 22.7 973.03 1
8 32 32 25.2 22.7 1714.29 1
8 32 35 na 22.7 1554.65 1
8 33 14 25.1 22.1 70.18 1.272382937
8 33 16 24.9 22.1 100.55 1.272382937
8 33 18 24.6 22.1 83.24 1.272382937
8 33 21 26.4 22.1 70.29 1.272382937
8 33 23 26.3 22.1 84.8 1.272382937
8 33 25 25.5 22.1 102.37 1.272382937
8 33 28 27 22.1 237.98 1.272382937
8 33 30 26 22.1 236.18 1.272382937
8 33 32 25.4 22.1 247.02 1.272382937
8 33 35 na 22.1 460.6 1.272382937
8 34 14 24.8 22.9 90.53 1.666759001
8 34 16 25.3 22.9 218 1.666759001
8 34 18 25.2 22.9 329.18 1.666759001
8 34 21 26.1 22.9 468.79 1.666759001
8 34 23 25.5 22.9 573.18 1.666759001
8 34 25 25.4 22.9 633.31 1.666759001
8 34 28 26.2 22.9 1240.04 1.666759001
```

```
8 34 30 25.6 22.9 1535.58 1.666759001
8 34 32 26.4 22.9 2419.09 1.666759001
8 34 35 na 22.9 2874.44 1.666759001
8 35 14 24.9 23.2 76.38 4.123728675
8 35 16 23.9 23.2 150.97 4.123728675
8 35 18 24.3 23.2 164.02 4.123728675
8 35 21 24.7 23.2 212.71 4.123728675
8 35 23 24 23.2 308.01 4.123728675
8 35 25 24.5 23.2 408.07 4.123728675
8 35 28 24.9 23.2 805.1 4.123728675
8 35 30 24.3 23.2 922.21 4.123728675
8 35 32 25 23.2 1345.59 4.123728675
8 35 35 na 23.2 1393.01 4.123728675
8 36 14 23.2 21.3 101.53 4.828372328
8 36 16 23.4 21.3 204.9 4.828372328
8 36 18 24.1 21.3 242.6 4.828372328
8 36 21 24.5 21.3 483.31 4.828372328
8 36 23 24.6 21.3 627.52 4.828372328
8 36 25 24.6 21.3 885.6 4.828372328
8 36 28 25.3 21.3 832.37 4.828372328
8 36 30 25.6 21.3 1765.5 4.828372328
8 36 32 25.4 21.3 2080.56 4.828372328
8 36 35 na 21.3 2350.28 4.828372328
```

Complete R code for this project

```
library(lme4)
## read data
mouse.data$ECM1_squaredxDay <- mouse.data$ECM1_level^2 * mouse.data$Day
mouse.data$ECM1_cubedxDay <- mouse.data$ECM1_level^3 * mouse.data$Day
mouse.data$logECM1xDay <- log(mouse.data$ECM1_level) * mouse.data$Day
mouse.data$logECM1xWeight_0xDay <- log(mouse.data$ECM1_level) *
mouse.data$Weight_0 * mouse.data$Day
## fit the Gaussian glm
gauss.m1 <- glm(formula = Tumor_volume ~ Day + ECM1xDay + Weight_0xDay,
family = gaussian(link = log), data = mouse.data)
summary(gauss.m1)
jpeg("gauss_m1_diagnostic_plot.jpg")
par(mfrow=c(2,2))
plot(gauss.m1)
dev.off()
## fit the gamma error glm
gamma.m3 <- glm(formula = Tumor_volume ~ Day + ECM1xDay + Weight_0xDay,
family = Gamma(link = 'log'), data = mouse.data)
summary(gamma.m3)
confint(gamma.m3)
jpeg("gamma_m3_diagnostic_plot.jpg")
par(mfrow=c(2,2))
plot(gamma.m3)
dev.off()
## model comparison
gamma.m1 <- glm(formula = Tumor_volume ~ Day, family = Gamma(link = 'log'),
data = mouse.data)
gamma.m2 <- glm(formula = Tumor_volume ~ Day + ECM1xDay,
family = Gamma(link = 'log'), data = mouse.data)
gamma.m4 <- glm(formula = Tumor_volume ~ Day + ECM1xDay + Weight_0xDay +
ECM1xWeight_0xDay, family = Gamma(link = 'log'), data = mouse.data)
anova(gamma.m1, gamma.m2, gamma.m3, gamma.m4, test = 'Chi')
## fit the Generalized mixed effect model
gamma.mm1 <- glmer(formula = Tumor_volume ~ Day + ECM1xDay + Weight_0xDay +
(1 + Day|Cage), family = Gamma(link = 'log'), data = mouse.data)
```