## Introduction and Aim

This report explores the capabilities of a new pharmaceutical treatment; primarily finding the maximum safe dosage of the treatment. Only 40 data points from a phase 1 clinical trial were available with three measured observations for one: the dose received, the measured response of the dose and the presence of a certain biomarker. The second aim of this report is to decide if the biomarker raises the tolerance of the patient in response to the treatment. The hypothesis test for this analysis is as follows: the null hypothesis is that the biomarker does not affect the tolerance of the patient to the treatment against the alternative hypothesis where the biomarker does affect the tolerance to the treatment. From our initial analysis of the data (shown below), we expect to accept the alternative hypothesis and be able to fit a model to the data as there appears to be a relationship between the dose and response that is distinct for those patients with and without the biomarker.

Dose response for population with/without biomarker
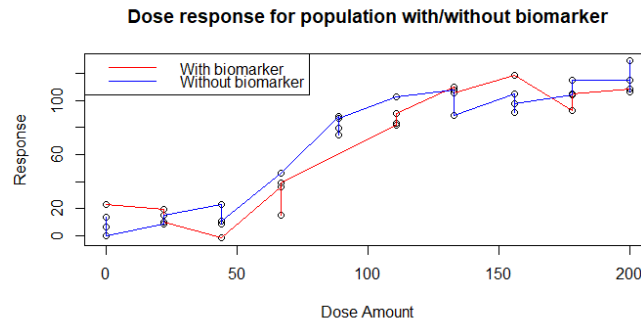
The model considered

To model the response of the dose, we use the four parameter E_{max} model:

$$r_i = E_0 + \frac{d_i^\lambda E_{max}}{d_i^\lambda + (ED_{50} + \beta x_i)^\lambda} + \epsilon_i$$

The terms are defined as follows: r_i is the measure response for the dose for the ith patient; d_i is the dose given to the ith patient; x_i indicates whether the ith patient has the biomarker (1 for present and 0 otherwise); E_0 is the response when the drug is not given; λ is the slope factor defined as sensitivity measured to the dose of the treatment; ED_{50} is the dose of the treatment that gives half the maximum response, E_{max}; β measures the change in ED_{50} with the biomarker is present in the patient; and ϵ_i is the error term for patient i given by $\epsilon_i \sim N(0,\sigma^2)$.

The E_{max} model is suitable for analysing this trial because: it adapts to the response without the drug present, denoted by E_0; it accounts for the presence of the biomarker, denoted by x_i, so it is possible to determine how and if the biomarker affects the patients' tolerance to the treatment, a key aim of the study; and the slope factor coefficient, λ, accounts for the sensitivity to the dose of the treatment and is helpful to understand the relative differences between the levels of dosage given.

Metropolis-Hastings: Results

To calculate the parameter values in the E_{max} model we used the Metropolis-Hastings algorithm. Our implementation used the function ll which returns the log likelihood of the model. It ensures that the only accepted proposals contain real numbers by returning -infinity if $ED_{50} + \beta \leq 0$. We also used the function pri which returns the logged prior of our parameters. It uses the following prior distributions for E_0, E_{max}, ED_{50}, λ, σ^2 and β respectively: Norm(0,10), Norm(100,10), Beta(2.5,5), Unif(0.5,3) and Norm(0,3) truncated with a lower bound of 0. This function also ensures that accepted proposals contain real numbers by returning -infinity if $ED_{50} + \beta \leq 0$.

We ran two versions of the Metropolis-Hastings algorithm. Firstly, we used a random walk proposal with starting values of 4.63, 115, 75, 3, 3.66 and 11.8 and standard deviations of 1.9, 2.4, 1.5, 0.04, 0.04 and 1.4 for E_0, E_{max}, ED_{50}, λ, σ^2 and β respectively. We then tested for correlation between the posterior distribution of each parameter in the model by calculating the sample correlation between each parameter and performing the Pearson's product-moment correlation test. The results (shown below) showed significant negative correlations between E_0 and E_{max} and between ED_{50} and β whilst E_0 and ED_{50} were significantly positively correlated. Next the Metropolis-Hastings algorithm was re-ran using the covariance matrix for the posterior distributions, the same starting values and a standard deviation of 0.9.
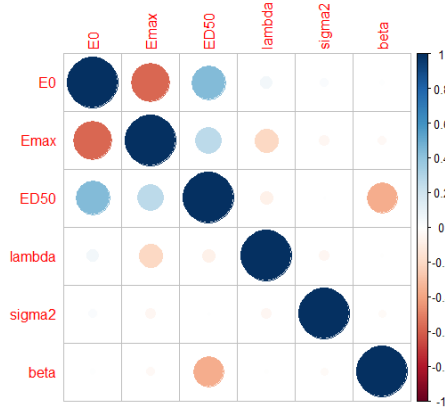
After running the Metropolis-Hastings both times we checked for convergence by subsampling independent observations of each parameter in both proposals. This independent subsampling is done by calculating the autocorrelation length, k_i, of the i-th parameter and keeping the k_i-th element of the vector of observed values for each parameter. We divided our independent samples into two populations and performed the two-sample Kolmogorov-Smirnov to test if the populations follow the same distribution. The tests showed that the samples of the posterior distribution obtained from both runs of the Metropolis-Hastings sampler converged as all parameters yielded a p-value greater than 5%. As both proposals were feasible we used the Deviance Information Criterion (DIC) to decide which had the best fit. The first sampler gave DIC of 127.946 whilst the second gave a DIC of 127.865. Although the difference is small, including the correlations from the posterior distribution provides a better fit and this is the model we shall use. The findings from both the Kolmogorov-Smirnoff tests and DIC match our expectations from running the algorithms as the trace plots showed good mixing and the cumulative average plots showed convergence. The plots shown below are only for the run including the correlations as the plots were alike for both runs as the similar DIC shows.
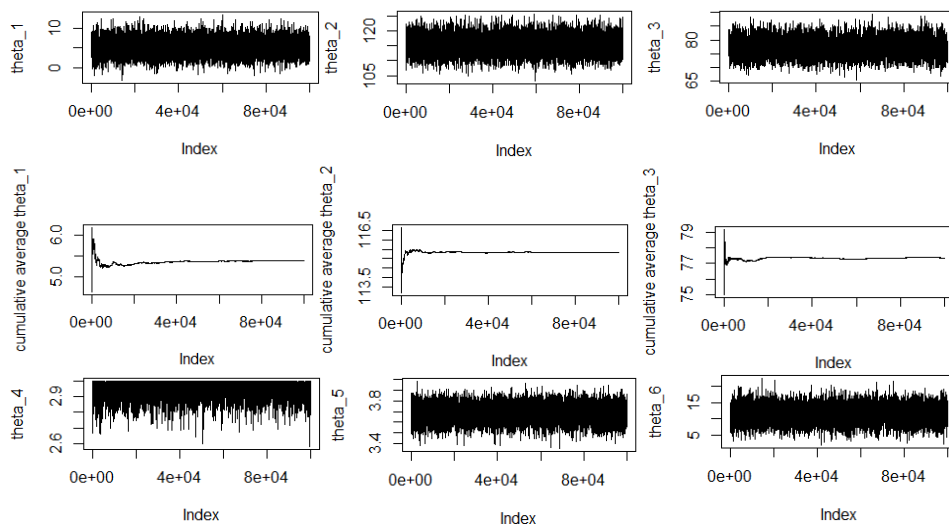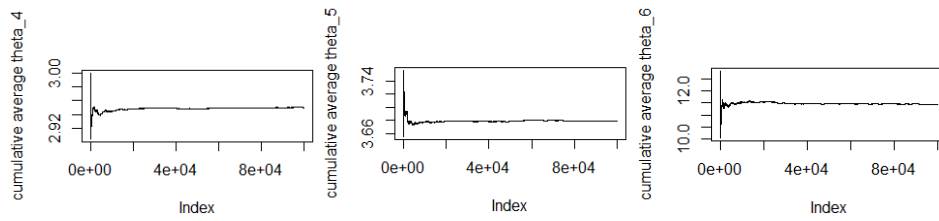
Model Checking

First, we verify that each parameter estimate is significant, i.e., different than zero. To determine this significance, we compute the 95% Credible Intervals (CI's) and check whether these intervals contain zero or not.

- 95% CI for E_0: 1.53-9.22
- 95% CI for E_{max}: 109.77-120.97
- 95% CI for ED_{50}: 71.76-83.24
- 95% CI for λ: 2.822-2.998
- 95% CI for σ^2: 3.524-3.825
- 95% CI for β: 6.56-16.23

Thus, we can conclude that each parameter is significantly different than zero. In particular, the parameter β, which is associated with a different response to the treatment in presence of the biomarker is different than zero.

The straight lines in the Q-Q plots below show that E_0, E_{max}, ED_{50}, σ^2 and β are approximately symmetrical about their means and the kernel density plots help to reiterate. Therefore, our model will take the mean values of these parameters which are 5.36, 115.38, 77.29, 3.68 and 11.42 respectively. However, the plots show that λ is significantly negatively skewed and hence we have decided to use the median value of λ, which is 2.97.
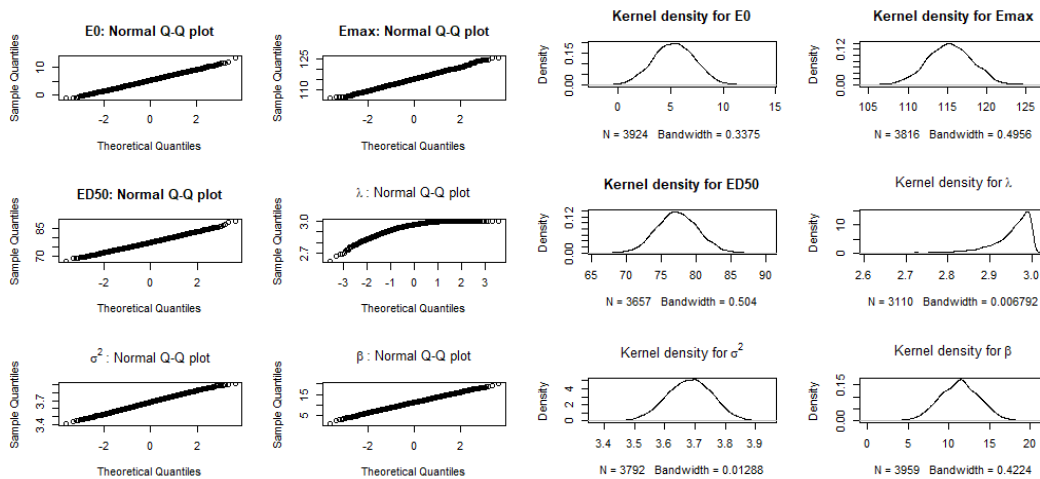
Using the estimated parameter values, we calculate the predicted response to treatment and the residuals as the difference between the predicted and the observed response. We can confirm that the presence of the biomarker provides a significant difference in the response to treatment (as shown in the figure below).

Predicted dose response
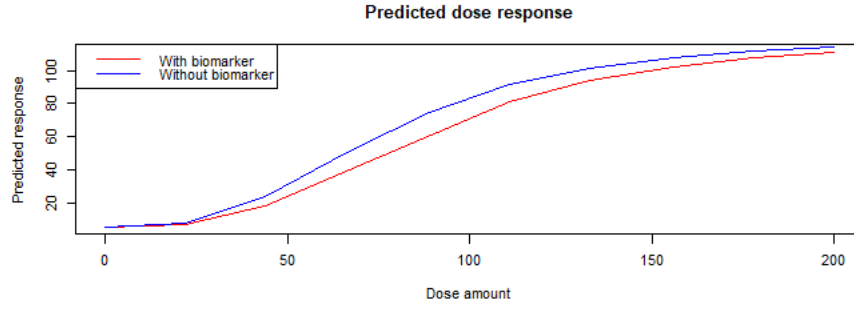
To assess if the model is providing a significant fit, we compute the residuals between the predicted and the observed responses to the treatment. From these plots we can observe that the residuals show no discernible pattern, providing evidence of them having constant variance. Further, we can see from the autocorrelation plot that there is no correlation between them. Finally, the normal quantile plot shows a linear trend, proving visual evidence that they are normally distributed, this was later confirmed by performing the Cramer-von Mises normality test. This allows us to conclude that the residuals come from a normal distribution and are uncorrelated, hence they are independent.
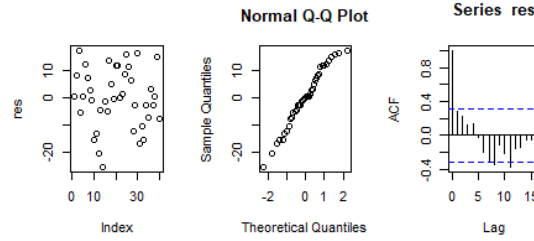
Conclusion

We are confident that our second Metropolis-Hastings sampler is accurate for predicting the E_{max} model. Using the estimated parameter values from this sampler, we have concluded that the dose response can be modelled by:

$$r = 5.3 + \frac{115.38 d^{2.97}}{d^{2.97} + (77.29 + 11.42x)^{2.97}} + N(0, 3.68)$$

As stated in the model checking section, β is a significant parameter in the model, so we can be clear that the biomarker increases the tolerance level of the treatment. Therefore, clinicians may be able to administer a higher dose to patients with the biomarker.

Having considered the tolerance levels of the individuals in the phase one of the trial with and without the biomarker, it would be advisable for the clinicians to set the maximal safe dose to ED_{30} for future trials. This is defined as:

$$ED_{30} = ED_{50} \left(\frac{0.3}{1 - 0.3} \right)^{\frac{1}{\lambda}}$$

Therefore, using our values for $\widehat{ED}_{50}$ and $\hat{\lambda}$, our value for ED_{30} is 58.1. This is the maximum dose the clinicians should use when giving the treatment to patients.

References

- Ting, N. (2006) *Dose Finding in Drug Development,* New York:Springer