Investigating the Toxicity of a New Pharmaceutical Treatment Using Bayesian Statistical Methods

Candidate numbers: 07297, 07291, 07276, 02286

Introduction

The aim of this report is to summarise the methods and analyses used to investigate the toxicity of a new pharmaceutical treatment, currently in a Phase I clinical trial. The clinicians were particularly interested in the dose that gave 30% of the maximum response to the treatment. Additionally, it was suggested by the clinical team that the presence of a biomarker in some patients may increase resistance to the treatment; this claim is also explored in the subsequent discussion.

Before using a Metropolis-Hastings (MH) sampler to estimate parameter values, we examined the structure of the data to gain a better understanding of the situation. A sample size of 40 patients were enlisted, 16 of whom were known to have the biomarker. They were randomly split into ten groups of four, with the patients in each group receiving the same dose amount- this ranged from 0 to 200, with increments of roughly 20 units. This resulted in a fair and random assignment of doses. The model that the following analysis is based on is the E_{max} model:
is there a more rigorous way of saying fair and random?

$$r_i = E_0 + \frac{d_i^{\lambda} E_{max}}{d_i^{\lambda} + (ED_{50} + \beta x_i)^{\lambda}} + \epsilon_i \qquad (1)$$

This model is described in detail in the next section. Of the parameters, E_0, E_{max}, ED_{50}, λ and β were estimated using Markov Chain Monte Carlo (MCMC) methods. Then we used these estimates to calculate an approximation of ED_{30} and found a safe maximum dose that the clinical team may use in the next stages of the trial.

Explanation of the Model

To investigate the dose-response relationship we considered the four-parameter E_{max} model defined in equation (1). The parameters and variables are as follows: i is a patient indicator ($i = 1, \ldots, 40$), d_i is the dose received, r_i is the measured response to the dose, x_i is an indicator of the presence of the biomarker for patient i, E_0 is the response in the absence of treatment, E_{max} is the maximum effect attributable to the treatment, E_{50} is the dose that gives 50% of the response E_{max} in the absence of the biomarker, λ is a slope factor showing the sensitivity of the response to the dose, β is the change in E_{50} in the presence of the biomarker and ϵ_i is the random error for patient i. We assume that ϵ_i are independent and identically distributed with $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$. We used the above model to predict the parameters $E_0, E_{max}, E_{50}, \lambda$ and β.

With that we estimated the dose giving 30% of the maximum effect E_{max}, the ED_{30}, as follows

$$ED_{30} = ED_{50}\left(\frac{0.3}{1 - 0.3}\right)^{1/\lambda}. \qquad (2)$$

This model requires that the dose-response relation is increasing, which we expect, and that response can be expressed as a continuous outcome. Then the E_{max} model dose-response curve is given by the expected value of the E_{max} model.

We also expect some sensitivity of the model curve to changes in parameters. In fact E_0 and E_{max} define lower and upper asymptotic values for the dose-response curve. Changes in E_0 affect the starting value of the curve and changes in E_{max} affect the range of the curve. Also the E_{50} is the dose that gives 50% of the response E_{max} in the absence of the biomarker, so a higher ED_{50} value indicates that a higher dose is needed to produce an effect. Similarly, the slope factor λ determines the steepness of the curve. The larger the value of λ, the more sensitive the response is to changes in the dose of the drug. [1]

Description of the Sampler

We performed Bayesian inference to get estimates of the parameters in the model from the observed data. Since the joint posteriors of the six random

variables were analytically intractable we sampled the posteriors using Markov Chain Monte Carlo (MCMC). [3]

Specifically, we used Metropolis-Hastings (MH) algorithms to simulate the posterior distribution θ with the given observed data. First of all, the initial value of each parameter's Markov chain was chosen to be the mean of each parameter's prior distribution. These are $E_0 = 5$, $E_{max} = 115$, $ED_{50} = 80$, $\lambda = 3$, $\sigma = e^{3.8}$, and $\beta = 8$. The proposal distribution for each parameter was at first normal, with mean given by the position of the current step and variance tuned until the Markov chain was mixing well and exhibiting an acceptance probability of approximately 24%. During the tuning, we ran chains for 100,000 steps, and found variances of 1.35 for E_0, 2 for E_{max}, 1.8 for ED_{50}, 0.03 for λ, 0.055 for σ^2, and 1.5 for β produced well mixing chains; which can be seen in figure (?). We also found that setting the initial value of each parameter's Markov chain to the mean of its prior resulted in a negligible burn-in time, which is also apparent from figure?
Include trace and ACF plots
Did we check for multimodality?

Posterior Distribution We used qqplot to check the posterior distribution. All the five parameters have a acceptable normal distribution except the parameter λ.

Updated MH sampler

We found significant correlation between parameters (shown in table 1), so replaced our first proposal distribution with a multivariate normal, with mean given by the current position in phase space and covariance matrix $\hat{\Sigma}$ proportional to the covariance matrix between parameters estimated from chains generated using the previous proposal. The probability density of moving from position θ_{i-1} to θ_i is therefore given by $\theta_i^{'} \sim \mathcal{N}(\theta_{i-1}, \hat{\Sigma}k^2)$ [2] where k is some constant of proportionality. We tuned the value of k to improve the convergence rate of the chain, finding $k = 0.9$ to be about optimal, which is consistent with the estimate of $k \approx \frac{2.4}{\sqrt{d}}$ [2] for high dimension d; which is 6 in this case. We proceeded to check for convergence by comparing the samples of the first and second halves of the chain using a Kolmogorov-Smirnov test, and found the null hypothesis (that the samples follow the same distribution) was not rejected; which we expect from a chain that has converged.
We then calculated the autocorrelation length of each parameter's chain and found them to range between 22.55 and 45.80, leading to effective sample sizes ranging from 2183.20 to 4434.19. This seemed consistent with what we saw displayed in both ACF and trace plots. We then created sub-sampled chains of length 2222 by taking every 45^{th} sample from each full chain, creating sub-sampled chains that we are satisfied comprise approximately independent samples from the posterior distribution of each parameter.

Results

For each parameter, E_0, E_{max}, ED_{50}, λ, β, and σ^2 we estimated the mean by taking the sample mean of the relevant subsampled chain, then found for each parameter a 95% credibility interval comprising all values larger than the smallest 2.5% of samples, and smaller than the largest 2.5% of samples. The mean and credibility interval for each parameter is listed in Table 1, along with the correlation between every pair of parameters.

	E_0	E_{max}	ED_{50}	λ	β	σ^2
mean	5.45	115	77.2	2.96	11.5	35.3
95% confidence interval	(1.73,9.05)	(110,120)	(72.0,82.3)	(2.85,3.00)	(6.85,16.2)	(30.0, 41.1)
correlation between						
E_0	1.00	-0.547	0.446	0.0431	0.0200	0.00671
E_{max}	-0.547	1.00	0.273	-0.184	-0.00836	-0.00239
ED_{50}	0.446	0.273	1.00	-0.0656	0.0133	-0.377
λ	0.0431	-0.184	-0.0656	1.00	-0.0784	0.0109
β	0.0200	-0.00836	0.0133	-0.0784	1.00	-0.0198
σ^2	0.00671	-0.0239	-0.377	0.0109	-0.0198	1.00

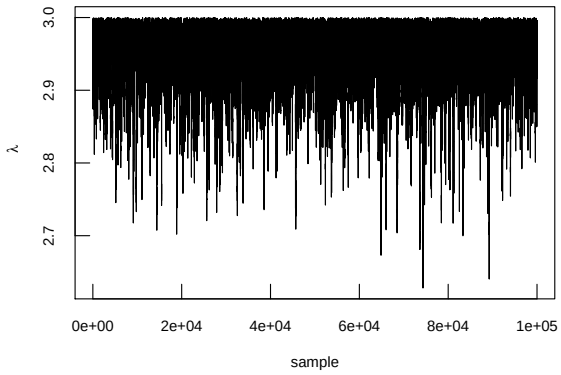
Table 1: Statistics of the posterior distribution of the E_{max} model (to three significant figures).

We were relieved to see that σ^2 is not significantly correlated with any other parameters, because it represents the normally distributed random variation between patients, which should not be related to the parameters of the model. Some correlation between the other parameters is present, but expected.

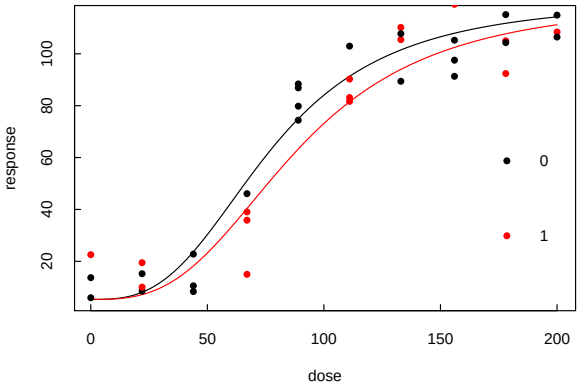
It may be that the prior distribution of the slope parameter λ is poorly chosen. Specifically, the prior strictly asserts that $\lambda < 3$, which appears (very unlikely). Look at this trace plot

The clinical team suspect that the presence of the biomarker increases a patient's tolerance to the treatment. To test this hypothesis, we noted first of all that $6.78 < \beta < 16.3$ with 95% credibility so β is positive with statistical significance. Then we looked at the E_{max} model and observed that, for positive β, the patient response r for a given dose d is lower if $x = 1$ than it would be if $x = 0$. Consequently, patients that exhibit the biomarker appear to be more tolerant to the treatment than patients who do not, which supports the clinical team's hypothesis. The modelled response of patients with or without the biomarker is compared to experimental data in figure .

A trace plot; look at how funny it is



The theoretical response to treatment as a function of dose for an arbitrary patient without the biomarker is represented by the black curve. This was calculated from the E_{max} model by taking the expectation of every parameter (listed in Table 1). The black theoretical response curve is compared to the black measured response points. Similarly, the theoretical response for patients with the biomarker is represented by the red curve and compared to the red dots representing the measured response. This plot does not clearly seperate the data points, something that may be rectified if λ were larger



Conclusion

Further to the effect of the biomarker, the clinical team is also interested in the dose causing a response with 30% the magnitude of the maximum response. We sought to estimate this by taking the subsampled chains of ED_{50} and λ, then defining a new chain with the same length, where the k^{th} entry is denoted ED_{30}^k and given by

$$ED_{30}^k = ED_{50}^k \left(\frac{0.3}{1 - 0.3} \right)^{1/\lambda^k} \tag{3}$$

where ED_{50}^k and λ^k are the k^{th} entries of the subsampled chains of ED_{50} and λ respectively. Next we took the mean of this chain, finding $\mathbb{E}(ED_{30}) = 64.3$ and constructed the 95% credibility interval $54.2 < ED_{30} < 62.0$ using the same two-tailed definition applied to other parameters in the model. It follows that $54.2 < ED_{30}$ with 97.5% credibility, so we advise the clinical team adopt a maximum safe dose of 54.2 for future trials.

References

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

[2] Wood, S. (2015). *Core Statistics*. Cambridge University Press

[3] Yildirim, I. (2012). *Bayesian Inference: Metropolis-Hastings Sampling.* [online] University of Rochester. Available at: http://www.mit.edu/ ilkery/papers/MetropolisHastingsSampling.pdf [Accessed: 7^{th} Dec. 2017]