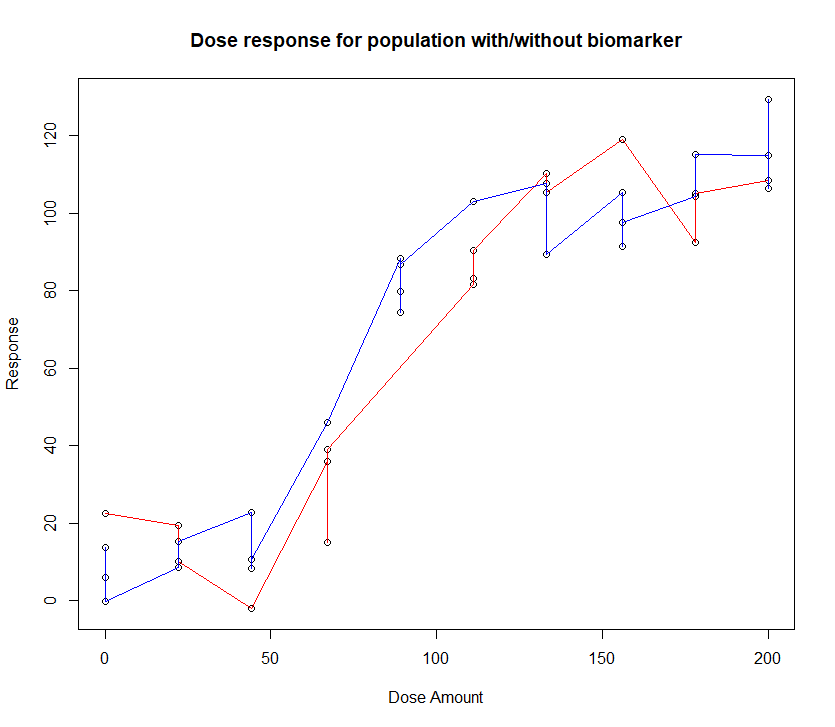
MA40198 Coursework

Introduction and Aim

This statistical report explores the capabilities of a new pharmaceutical treatment; primarily finding the maximum safe dosage of the treatment. The collated data is from a Phase one clinical trial where only a few patients (40 in total) are tested on and, if deemed safe, the treatment will pass onto the next stage of the clinical trial. Three observations were measured in each of the patients: the dose received, the measured response of the dose and the presence of a certain biomarker.

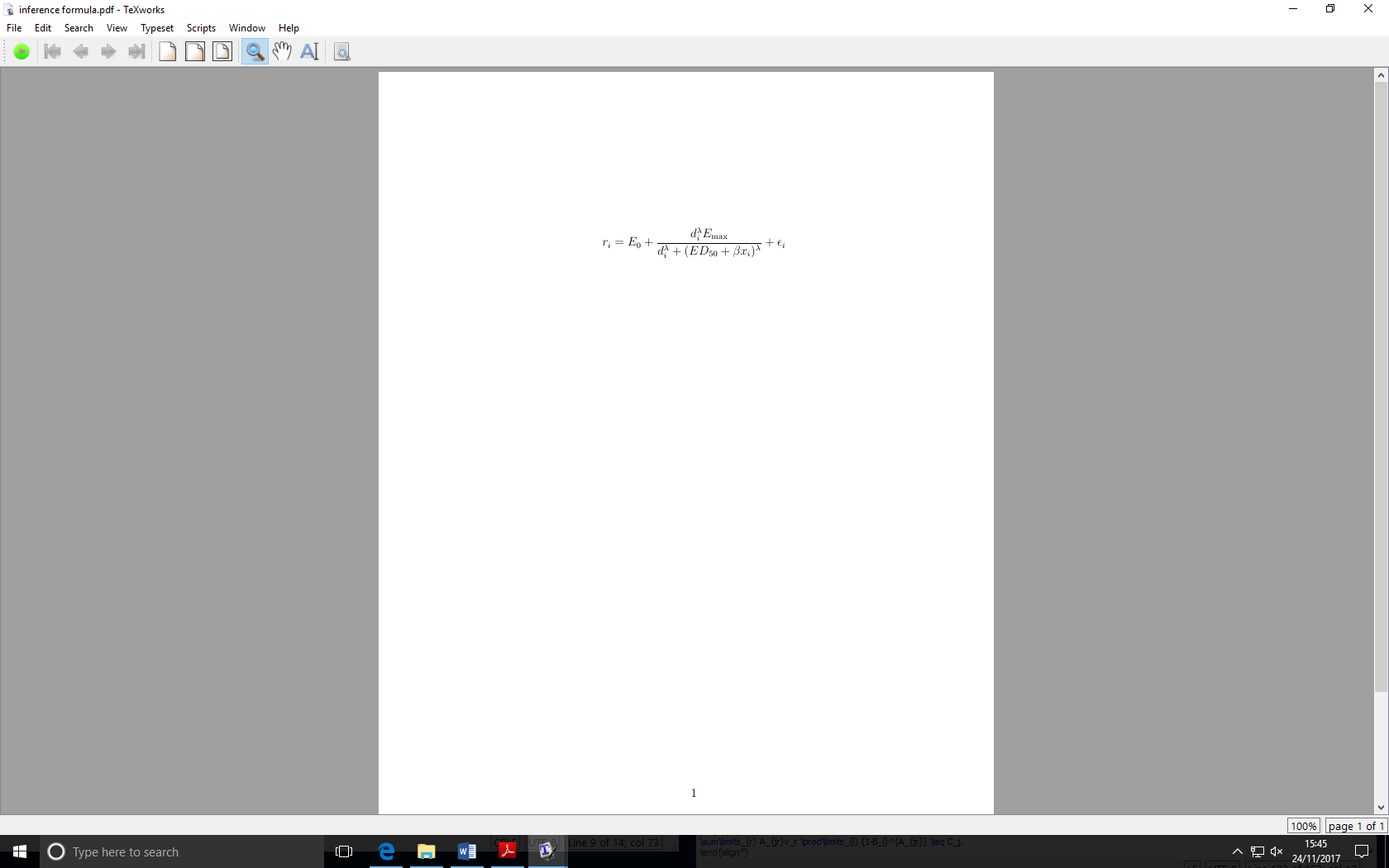
The second part of this analysis is to study whether the biomarker, identified by clinicians, will raise the tolerance of the patient in response to the treatment. Therefore, the hypothesis test for this analysis is as follows: the null hypothesis is that the biomarker does not affect the tolerance of the patient to the treatment against the alternative hypothesis whereas the biomarker does affect the tolerance to the treatment. However from our initial analysis of the data, we plotted a graph (shown below) of the response of the individuals against the dose amount with the biomarker (in red) and without the biomarker (in blue). From this graph we expect to accept the alternative hypothesis and be able to fit a model to the data



Finally, the third part of this analysis is to know more about giving 30 percent of the maximum dosage in order to set the maximum safe dosage for the treatment; denoted as ED30 here and defined in full below where the clinicians want to use the highest available dose below ED30.

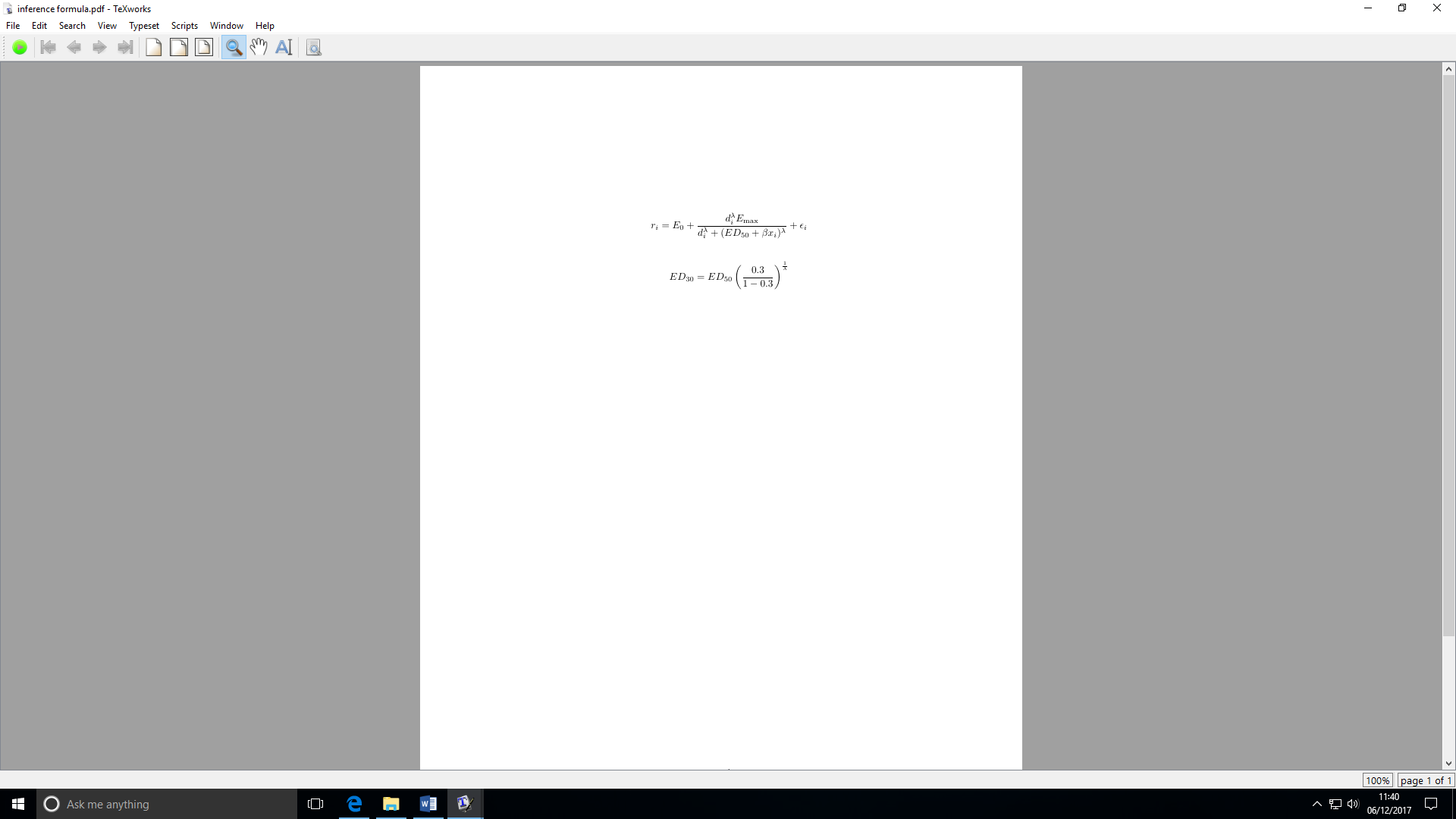
The model considered

To model the response of the dose, we use the four parameter Emax model. This model is shown below:



Where the terms can be defined as follows: ri := is the measure response for the dose for the ith patient; di := is the dose given to the ith patient; xi:= indicates whether the ith patient has the biomarker (given as 1) or not (given as 0); E0:= this is the response when the drug is not given, sometimes given as the placebo effect the patient has when given the treatment in which the model adjusts for this; λ:= is known as the slope factor defined as sensitivity measured to the dose of the treatment; ED50:= the dose of the treatment which is half of Emax; β:= measures the change in ED50  with the biomarker is present in the patient; εi:= is the error term for patient ith for εi~N(0,σ2).

Furthermore, ED30 is defined as:



The Emax model can be deemed suitable for analysing this trial in several ways. Firstly, this model adapts to the placebo effect, E0, therefore the model shifts around E0 meaning that this will not lead to a distortion in the results and giving more credit to the treatment. Secondly, the model accounts for the presence of the biomarker, denoted by xi, so it is possible to determine how and if the biomarker affects the patients’ tolerance to the treatment; which is one of the aims if this analysis. Thirdly, the use of 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.

The function ll returns the log likelihood of the four parameter Emax model. An output is only calculated if ED50+beta1 is greater than zero else the output is minus infinity. This ensures that we are only using real numbers and positive doses. In physical terms, the patient would have an unlimited tolerance to the treatment which is not possible. Therefore the function would just return out which is defined as negative infinite.

The function pri returns the prior of our parameters where E0 has a normal distribution with mean 0 and standard deviation 10. Emax also has a normal distribution with mean 100 and standard deviation 10. ED50 has a beta distribution with shape parameters 2.5 and 5. Lambda has a uniform distribution between the interval 0.5 and 3. The standard deviation has a proposed truncated normal distribution with mean 3 and standard deviation 5. Finally beta has a normal distribution with mean 10 and standard deviation 4. Then we are returning the log likelihood of our proposed distribution making sure ED50+beta1 is greater than zero.

Metropolis-Hastings: Results

To calculate the parameter values in the Emax model we used 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 Eo, Emax, ED50, λ, σ2 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 E0 and Emax and between ED50 and β whilst E0 and ED50 were significantly negatively 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.

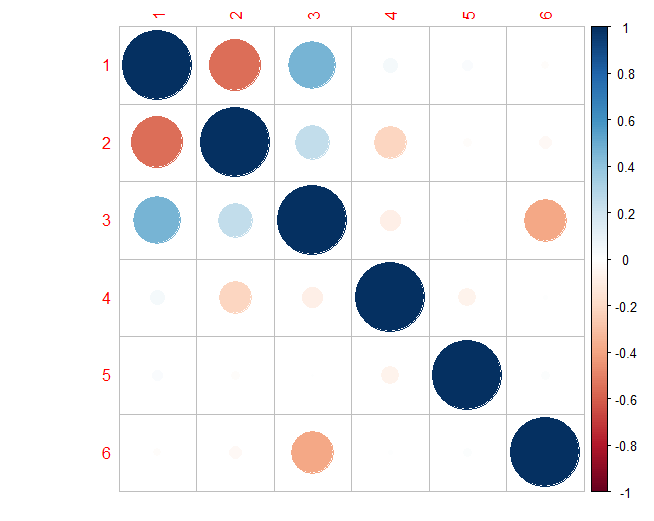
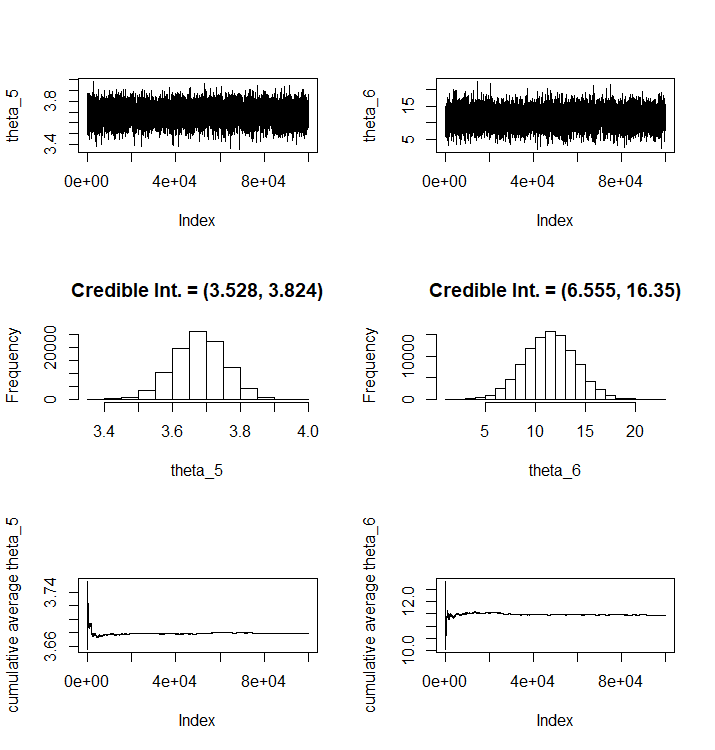
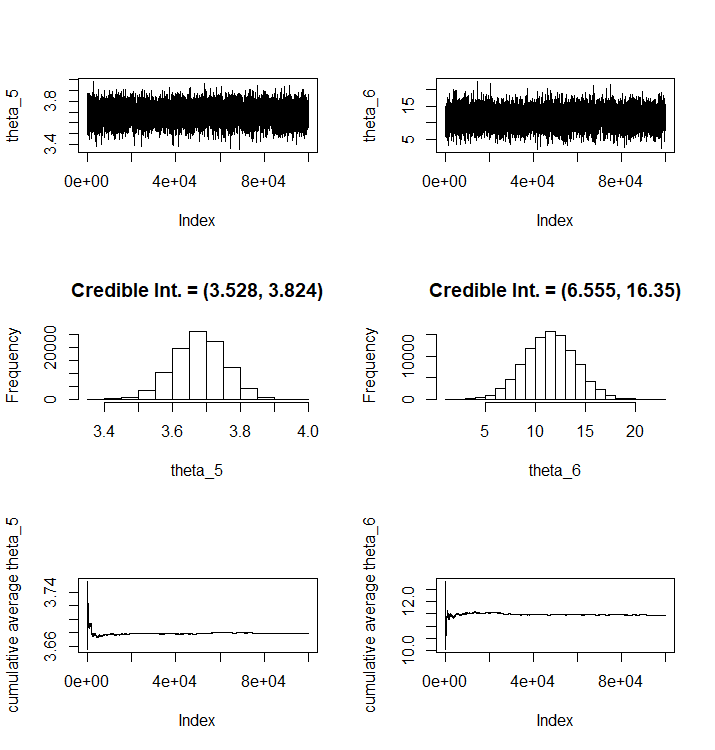
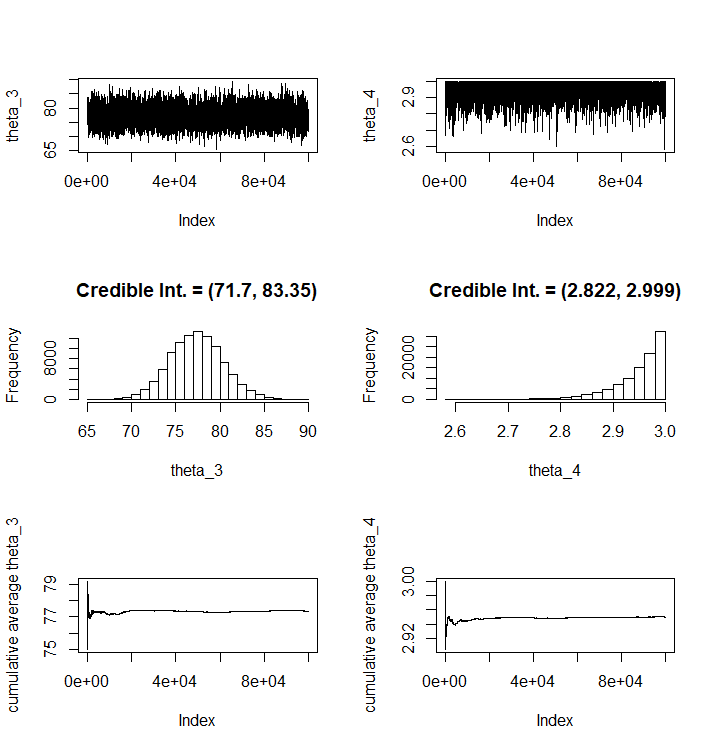
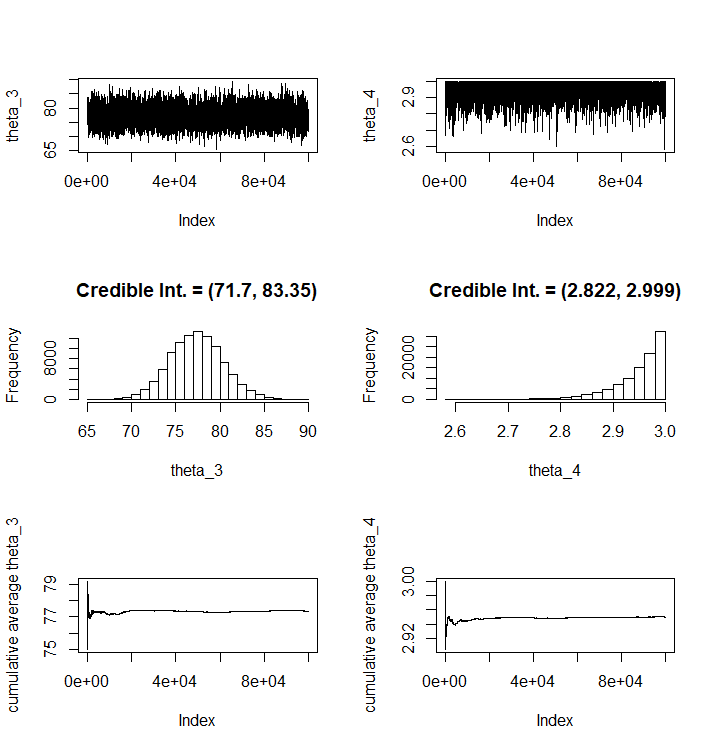
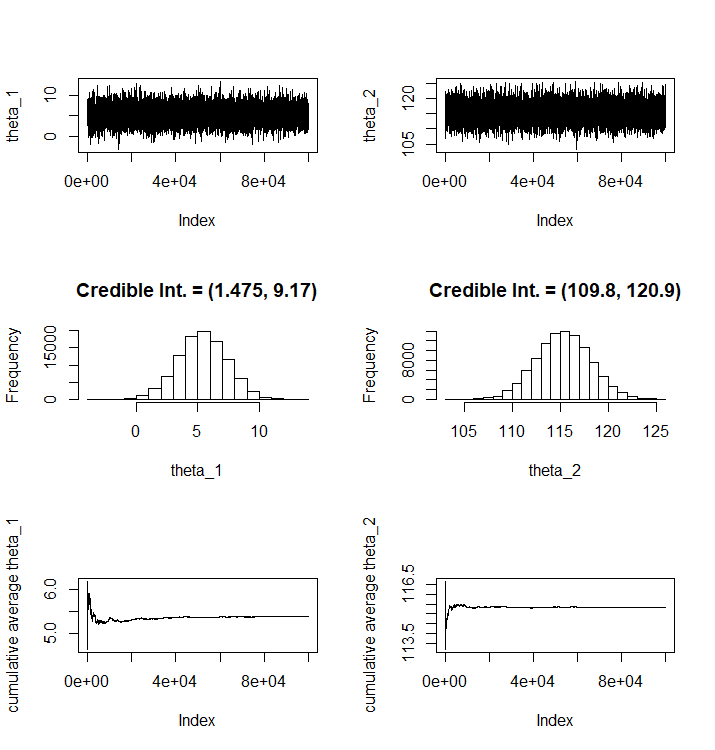
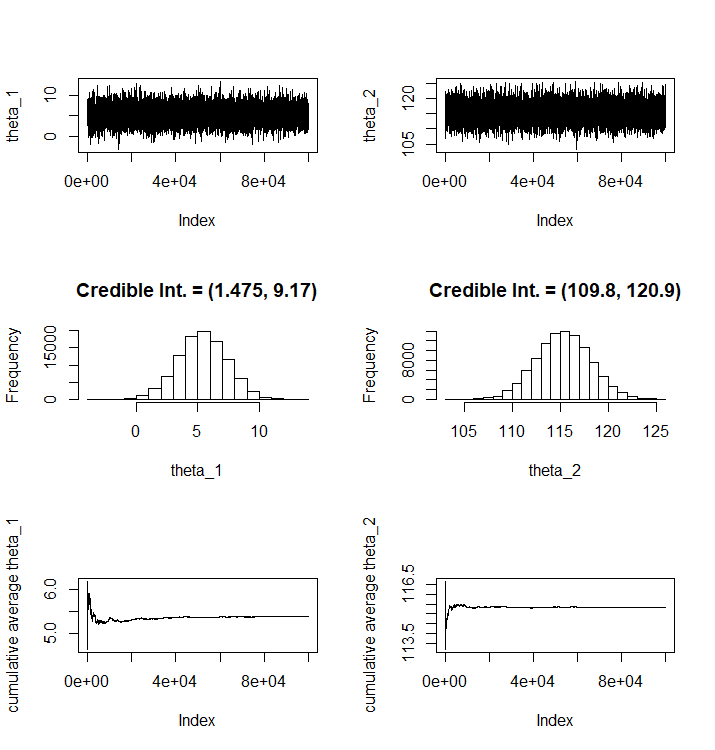


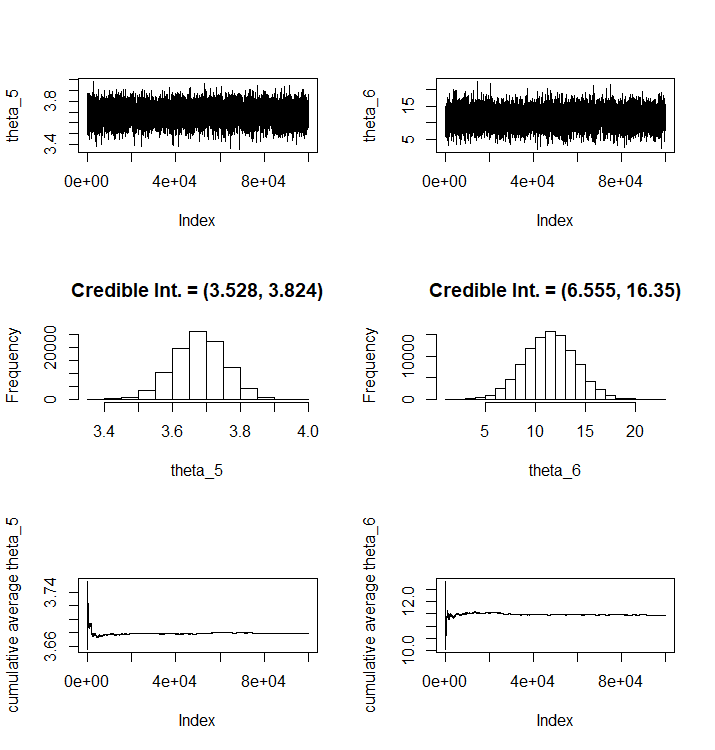
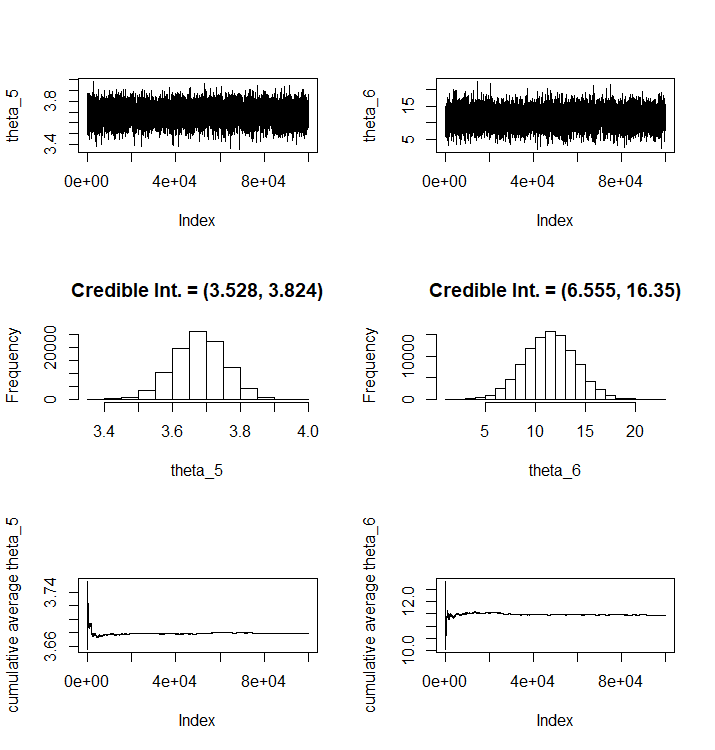
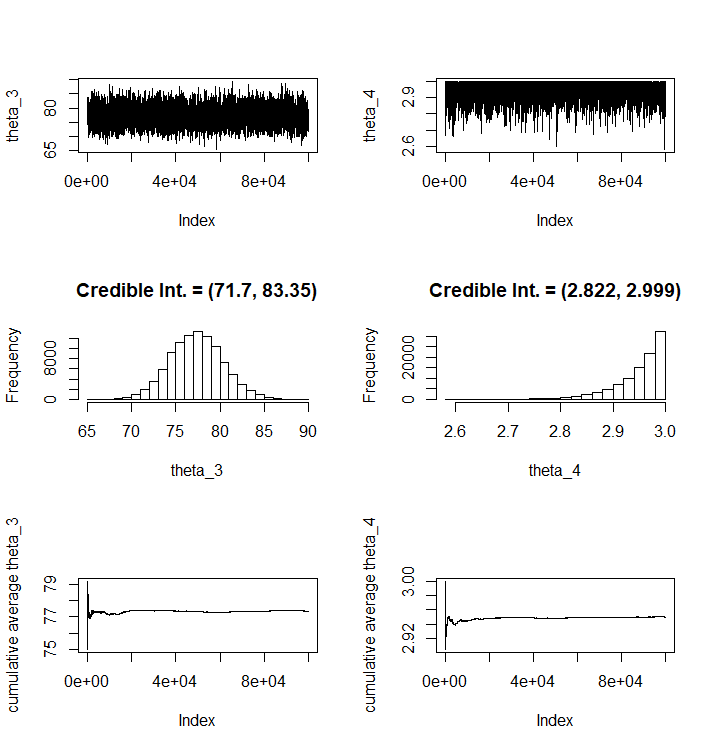
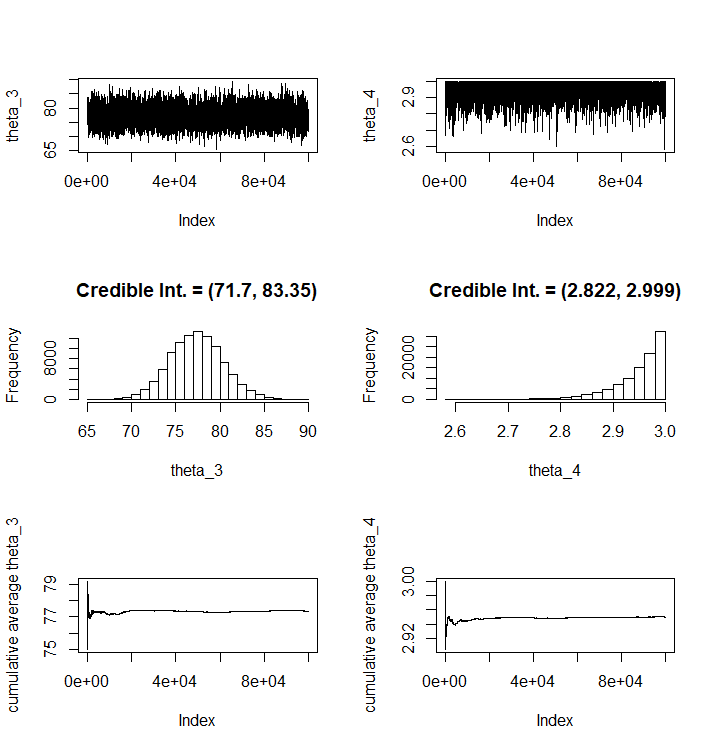
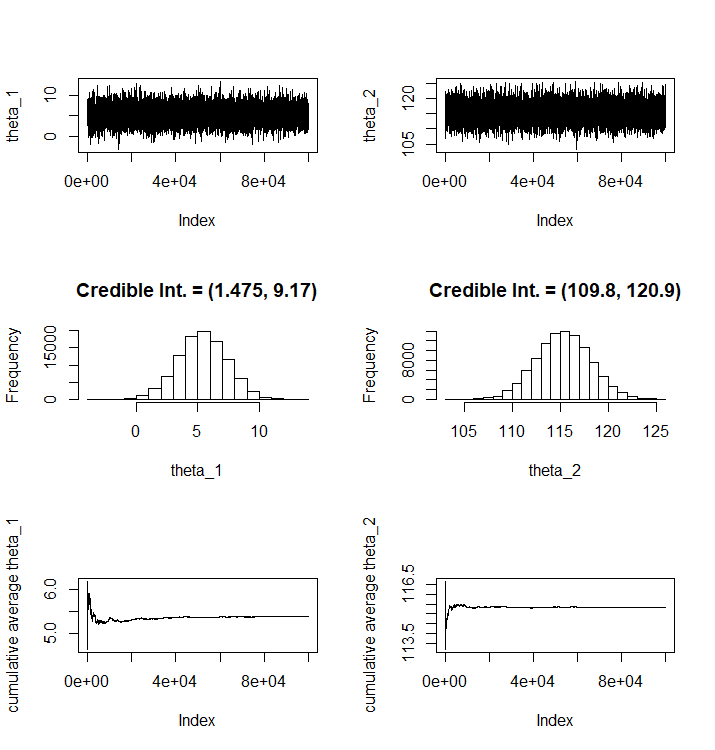
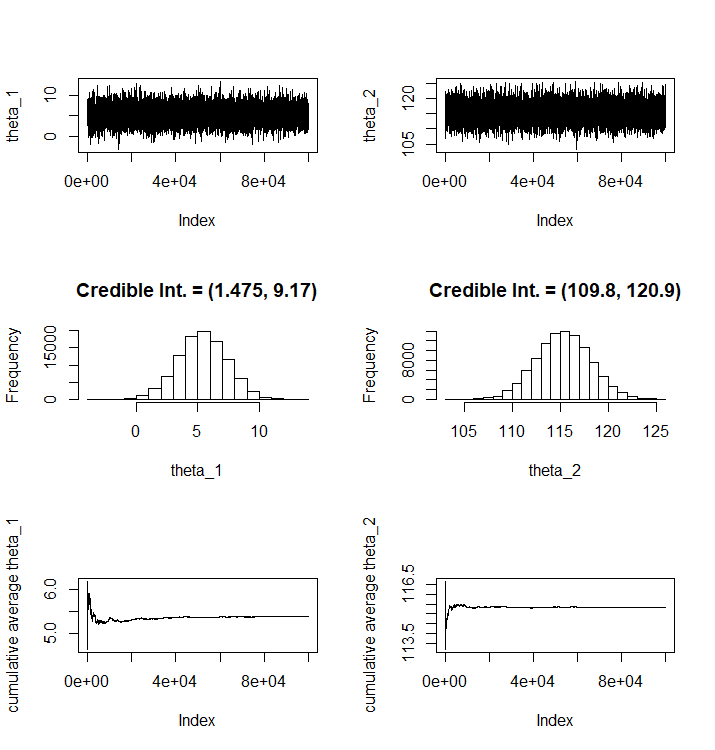
Figure 2: A pictorial diagram of the Pearson’s product-moment correlation test between the six parameters

After running the Metropolis-Hastings both times we checked for convergence by subsampling independent observations of each parameter, for each of the two proposals. This independent subsampling is done by calculating the autocorrelation length, , of the -th parameter and keeping the -th element of the vector of observed values for each parameter. Afterwards, we divide our independent samples into two populations and perform the two-sample Kolmogorov-Smirnov to test if the populations follow the same distribution. We say that the proposal for the posterior distribution converged if, for each parameter, the Kolmogorov-Smirnoff test yielded a p-value greater than 5%. The Kolmogorov-Smirnoff tests showed that the samples of the posterior distribution obtained from both runs of the Metropolis-Hastings sampler converged.

As both proposals were feasible we decided which was the better fit by calculating the Deviance Information Criterion (DIC) for each. The first sampler gave DIC of 127.9149 whilst the second gave a DIC of 127.7179. Even though 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 Deviance Information Criterions match our expectations from running the algorithms. The trace plots for both runs were very similar so only the trace plots corresponding to the second run are shown. The trace plots show good mixing and the cumulative average plots show convergence.





Conclusion

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 ED30 for future trials.

<http://www.math.chalmers.se/~rootzen/finrisk/reportwriting0315.pdf> : for how to write a report

<https://www.wikihow.com/Write-a-Statistical-Report>

<http://file.zums.ac.ir/ebook/75-Dose%20Finding%20in%20Drug%20Development%20(Statistics%20for%20Biology%20and%20Health)-Naitee%20Ting-0387290745-Sp.pdf>