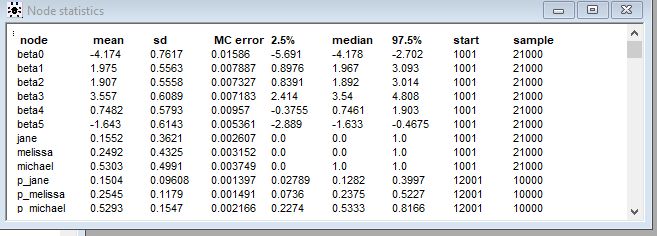
**Kien Duc Vu Final Exam**

**Problem 1:**

-95% credible set of probability of Jane is [0.02789, 0.3997] with mean = 0.1504. The probability of Janes going to the beach calculated in homework 1 is 0.17238, which is fell in 95% credible set.

-95% credible set of probability of Michael is [0.2274, 0.8166] which mean = 0.5293. The probability of Michael going to the beach calculated in homework 1 is 0.40704, which is fell into 95% credible set.

-95% credible set of probability of Melissa is [0.0736, 0.5227] which mean = 0.2545. The probability of Melissa going to the beach calculated in homework 1 is 0.27964 which is fell into 95% credible set.



**Problem 2:**

1. 95 % credible set of LC50 predicted from provided values is [1.877, 6.617] with mean = 4.234.

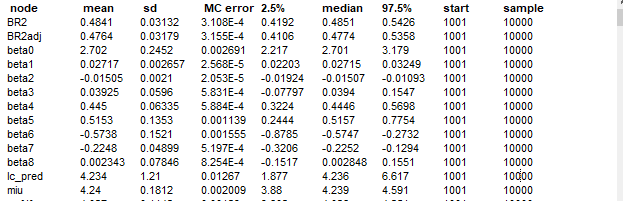
And the 95% credible set of mean of LC50 is [3.88, 4.591] which bar(miu) is 4.24.

1. H-050 and C-040 can be ignored while fitting Bayesian multilinear regression since their 95% credible set contain 0 and their means ~ 0.

H-050 95% credible set = [-0.07797, 0.1547] with mean = 0.03925

C-040 95% credible set = [-0.1517, 0.1551] with mean = 0.002343

1. Bayesian R-square = 0.4841 with 95% credible set [0.4192, 0.5426]. Since Bayesian R^2 is centered around 0.4841, the provided features to predict LC50 can only explain 50% variance in the prediction. Some of the predictors must be missing.

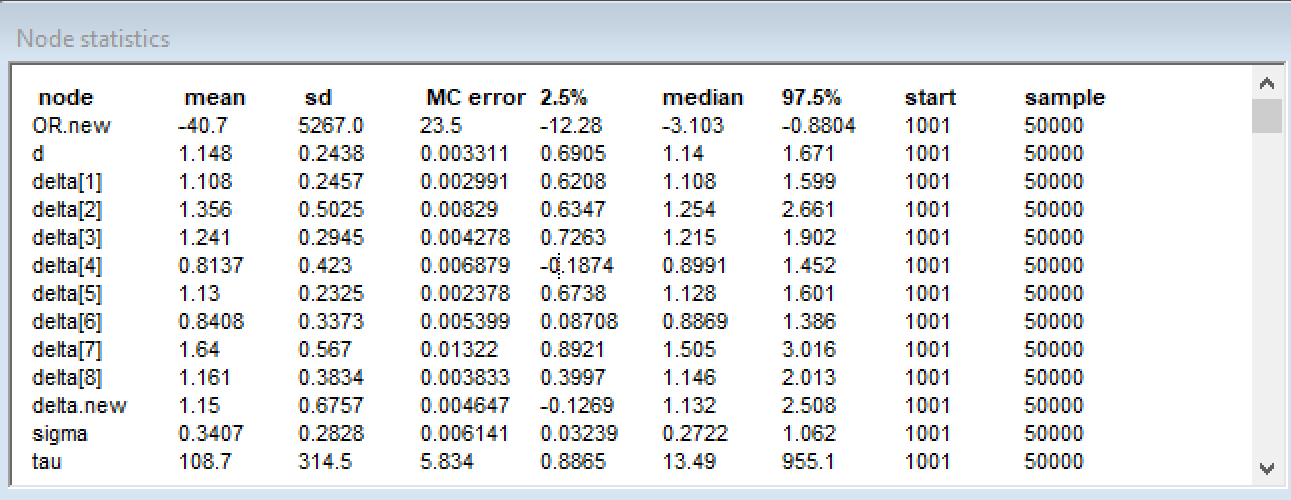


**Problem 3:**

1. We might believe that it is unreasonable to assume that all the studies in our meta-analysis are estimating exactly the same treatment effect, and that are they are the only studies in which we are interested, or perhaps the only studies that exist on the topic. We may assume that differences in samples, design, and conduct introduce more statistical heterogeneity between studies, than can reasonably be attributed to random error within studies. In that case it might be more reasonable to use a random effects model. A random effects model assumes the studies are a sample from all possible studies and includes an additional variance component for variation or heterogeneity between studies.

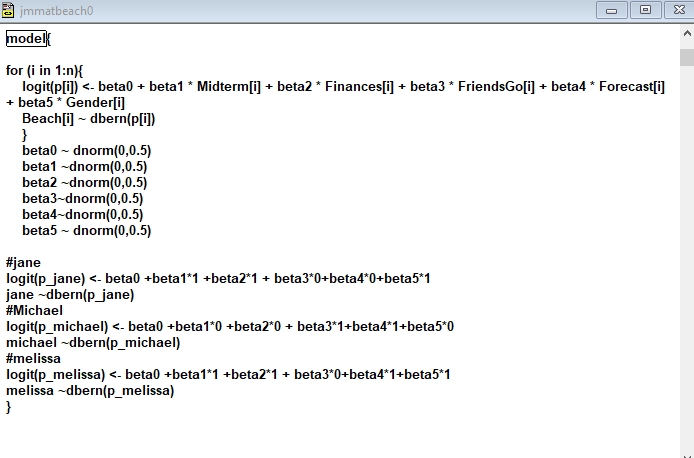
In this case, it is binomial distribution, where: (please see appendix for model)

* pc[i], pt[i] are the total number of patients in the two arms of the study, or the drug/placebo populations
* rc[i], rt[i] are the number of events in the two arms
* pc[i], pt[i] are the underlying probabilities used to define the likelihood. This is used to model in-between heterogeneity
* mu[i] is the log odds event in group A (and requires a prior)
* delta[i] is the so-called “treatment effect” or the log OR for the ith study, and
* d is the study-level point estimate, with tau the between study variance
* delta.new is the predictive probability distribution for most likely outcome in a future study

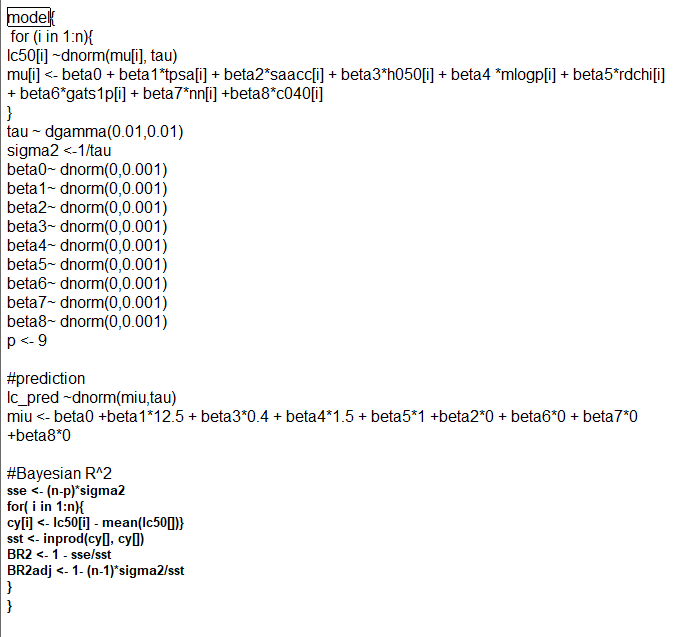


Based on Bayesian meta-analysis accounting for in-between heterogeneity, drug treatment from 8 studies is most likely beneficial with 95% credible drug effect is [0.6905, 1.671] with mean of 1.14. Future study predictive drug treatment effect is most likely beneficial also, which 95% credible set of predictive future effect is [-0.1269, 2.508] with mean of 1.132.

**Problem 1: Model**



Problem 2 Model:



**Problem 3-Model**

