

Statistical Inference Final Project

1. Introduction

1.1 General Idea of Meta-Analysis

Meta-Analysis is a method for systematically combining both pertinent qualitative and quantitative study data from several studies to develop a single conclusion that has greater statistical power. These selected studies may have different conclusions. Meta-analysis's conclusion is statistically stronger than the analysis of any single study, due to increased numbers of subjects, greater diversity among subjects, or accumulated effects and results. Generally it can give the study data a greater statistical power and confirmatory data analysis. Also, it can have greater ability to extrapolate to general population affected. In a word, meta-analysis pools together the populations from different studies into one statistical analysis and treats them as one large study population with one conclusion ^[1]. Normally study results could vary from one study to another. Using meta-analysis can give confirmatory study analysis and come out with only one single result. For example, 8 studies of streptokinase suggested its effectiveness in treating patients presenting with myocardial infarction, yet only 3 of these studies reported statistically significant results. The results of a meta-analysis combining data across all 8 studies concluded that streptokinase was associated with a statistically significant reduction in mortality ^[2].

In meta-analysis, the number of trials matter as well as the quality of the trial included. The techniques used in meta-analysis provide a structured and standardized approach for analyzing prior findings in a specific topic in the literature. Meta-analysis findings may not only be quantitative but also may be qualitative and reveal the biases, strengths, and weaknesses of existing studies. The result could be misleading if the meta-analysis only include poor quality studies. It is also important to account for varying sample sizes across trials. Usually sample size varies greatly among trials. The result of the meta-analysis is based on the weighted average from the results of selected studies.

Meta-analysis may suffer from publication bias because studies with negative results are less likely to be published. This can impact the selection of studies for a meta-analysis. Therefore, results from meta-analysis may overstate the positive results.

1.2 Differences between Bayesian and Frequentist

The Bayesians model uncertainty by a probability distribution over hypotheses. One's ability to make inferences depends on one's degree of confidence in the chosen prior, and the robustness of the findings to alternate prior distributions may be relevant and important. The frequentist school only uses conditional distributions of data given specific hypotheses. The presumption is that some hypothesis is true and that the observed data is sampled from that distribution. In particular, the frequentist approach does not depend on a subjective prior that may vary from one investigator to another [4]. For example, Bayesian and frequentist have different views in interpreting the 95% confidence interval. For Bayesian, they would interpret it as that there is a 95% probability that the parameter is in this interval. But for Frequentists, they would say that there is 95% confidence that the parameter is between the intervals [5].

1.3 Relationship between Meta-Analysis and Bayesian

As the Higgins and Spiegelhalter indicated in their study, the early trial results of magnesium have been said to be 'too good to be true'. Bayesian approach was proposed to formally accommodate the interpretation of usually strong treatment effects on the meta-analysis and meta-trials. In this way, they can emphasize the skepticism. In this study, they use a Bayesian perspective to possibly influence the interpretation of the published evidence on intravenous magnesium in acute Myocardial Infarction. Bayesian analysis will allow them to incorporate the skepticism in this study ^[6].

1.4 Controversy of the Use of Intravenous Magnesium in Patients with Acute Myocardial Infarction

Intravenous magnesium has long been believed to play an important role in patients with acute myocardial infarction. But there have been many responses to the apparent contradiction between the meta-analysis and the meta-trial. First, Yusuf responded to the ISIS-4 results and Flather pointed out that "since most treatments product either no effect or at least moderate effects on major outcomes such as mortality, investigators should be skeptical if the results obtained deviate substantially from these expectations. Some studies also showed that the choice of statistical methods used for the meta-analysis can generate substantially different results. Some researchers also claimed that there were some limitations in the meta-trials. For example, Borzak and Ridker pointed out that "an important limitation of the ISIS-4 protocol was that the actual time of magnesium initiation was unknown".^[6]

2. Introduce the Models

2.1 The Difference between Fixed Effects and Random Effects Meta-Analysis

Fixed effect assumes there is only one true effect size, which underlies all studies in the meta-analysis. The fixed effects model assumes that all included studies investigate the same population, use the same variable and outcome definitions, etc. The differences among studies are purely random errors or noises. We can express a simple fixed effect model as below:

$$= \beta + \varepsilon$$

In this model, we can view β as the true treatment effect and ε_i as the random noise, while $\varepsilon \sim N(0, \sigma^2)$. The fixed effect model provides a weighted average of a series of study estimates. The inverse of the estimates' variance is commonly used as study weight, such that larger studies tend to contribute more than smaller studies to the weighted average. Fixed effect model only considered the within study variance. So when we calculate weight, we can express it as

$$w = \frac{1}{\sigma^2}$$

Random effects emphasized heterogeneity in the data. It treats the variances as two types- within study variance and between study variance. So in this random effect model, the model's variance is expressed with two types:

$$= \beta + \varepsilon + \varphi$$

$\varepsilon_i \sim N(0, \sigma_i^2)$, and $\varphi_i \sim N(0, \sigma_\tau^2)$, and they are independent of one another. Now the weighted average can be expressed as

$$w_i = \frac{1}{\sigma_i^2 + \sigma_\tau^2}$$

In random effect, they assume studies are different, and there is a distribution of study effects. Heterogeneity is considered and is the key to separate with fixed effect models. If there is evidence of heterogeneity among the population effects, then random effect models should be considered. Under the fixed effect model, all studies are estimating the same effect size, and so we can assign weights to all studies based entirely on the amount of information captured by that study. A larger study could give more weight on it. By contrast, random effect model is trying to estimate the mean of a distribution of true effect. Large studies may yield more precise estimates than small studies, but each study is estimating a different effect size, and each of these effect size serve as a sample from the population whose mean we want to estimate [7-8].

2.2 Dataset

The dataset provided contains 15 different trials. It also provides the year of the trials. We can see the trials were performed during 1980s and 1990s. The dataset also includes both treatment groups and control groups. For each group, they also have the number of patients and deaths. In the ISIS-4 study, we can see there are 29011 patients in the treatment groups and 29039 patients in the control groups. They are substantially larger than other studies. In the fixed effect model, ISIS-4 may have bigger impact on the meta-analysis results than the random effect model. The details of the dataset can be found in table 2 on the below.

2.3 Peto Method Model Specifications

The Peto method is a widely used statistical method in fixed effect model. It can be used to summarize the odds ratio. Normally case-control studies of dichotomous (binary) outcomes can be arranged by a 2*2 table. In this meta-analysis's case, the table would be arranged by the below:

	Treatment	Control
Dead	a	b
Alive	c	d

Here we have n number of sample size, where $n=a+b+c+d$. k is the number of trials.

Peto odds ratio can be calculated by using this [9]:

$$\varphi = \exp\left(\frac{\sum_{i=1}^k (O_i - E_i)}{\sum_{i=1}^k V_i}\right)$$

Where

$$O = a$$

$$E = (a + b)(a + c)/n$$

$$V = \frac{(a + b)(c + d)(a + c)(b + d)}{n^2(n - 1)}$$

Peto odds ratio method has become the very fundamental choice for binary studies. However, it could lead to biased estimates for the odds ratio when the treatment effects are large or the group size ratio is not balanced ^[10].

2.4 DerSimonian and Laird Method Model Specification

DerSimonian and Laird (DL) method is commonly used method in meta-analysis random effect model. It initially assumes the estimate of treatment effect from the i^{th} study, Y_i is distributed as $Y_i|\mu_i \sim N(\mu_i, \sigma_i^2)$, where μ_i is the true underlying treatment effect of the i^{th} study and σ_i^2 is the corresponding within study variance. In DL, it also assume that $\mu_i \sim N(\mu, \tau^2)$, where μ is the overall treatment effect and τ^2 is the between- study variance. They are also independent of each other ^[11].

In DL method model, Q statistics is applied.

$$Q = \sum_{i=1}^n w_i (y_i - \bar{y})^2$$

Where $w_i = \sigma_i^{-2}$, $\bar{y} = \sum_{i=1}^n w_i y_i / \sum_{i=1}^n w_i$. Similar to Peto method, n is the number of trials/studies. Then we can calculate the expectation of Q:

$$E(Q) = (n - 1) + (S_1 - \frac{S_2}{S_1})\tau^2$$

Where $S_r = \sum_{i=1}^n w_i^r$, which provides the DerSimonian and Laird estimate

$$\tau_{DL}^2 = \max(0, \frac{Q - (n - 1)}{S_1 - \frac{S_2}{S_1}})$$

Then the estimate for the treatment effect would be

$$\hat{\mu}_{DL} = \frac{\sum_{i=1}^n \frac{y_i}{\sigma_i^2 + \tau_{DL}^2}}{\sum_{i=1}^n \frac{1}{\sigma_i^2 + \tau_{DL}^2}}$$

This method was the simplest method for random effect model. This is a relatively more sophisticated methodology when the sample size is large or the group size is not balanced.

3 The Frequentist Analysis

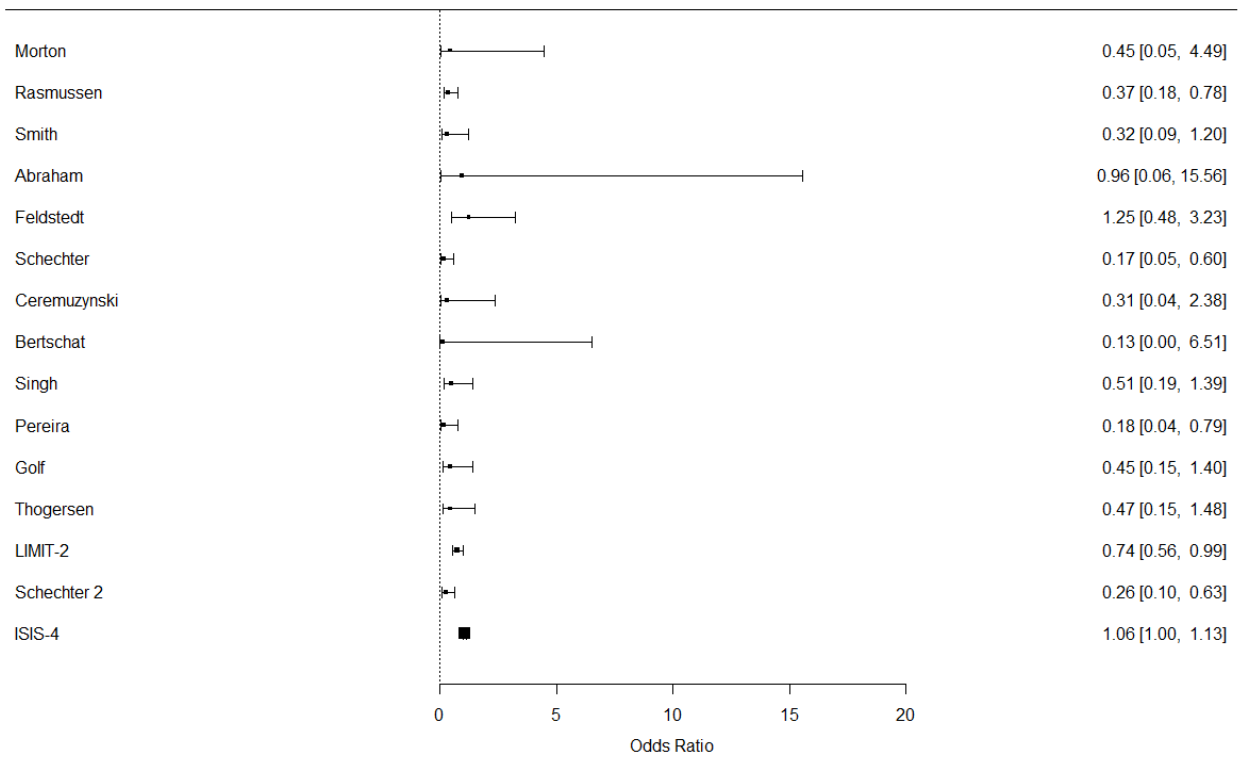
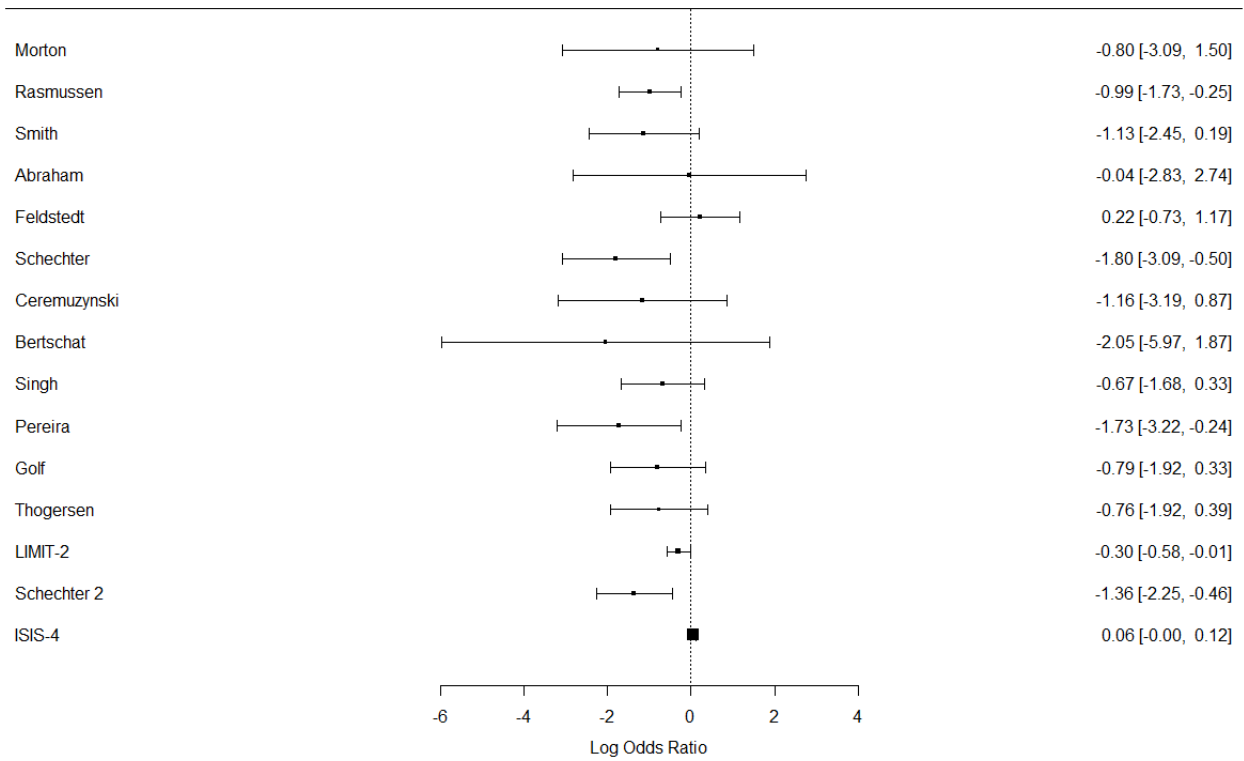
3.1 Reproduce the Table 2

Trial	Magnesium group		Control group	
	Deaths	Patients	Deaths	Patients
Morton	1	40	2	36
Rasmussen	9	135	23	135
Smith	2	200	7	200
Abraham	1	48	1	46
Feldstedt	10	150	8	148
Shechter 1990	1	59	9	56
Ceremuzynski	1	25	3	23
LIMIT-2	90	1159	118	1157
Fixed Effect (Peto) meta-analysis of above 9 trials: OR =0.65(95%CI: 0.51,0.82)				
Random Effect (DL) meta-analysis of above 8 trials: OR=0.55 (95% CI: 0.34,0.89)				
Bertschat	0	22	1	21
Singh	6	76	11	75
Pereira	1	27	7	27
Golf	5	23	13	33
Thogersen	4	130	8	122
Shechter 1995	4	107	17	108
Fixed Effect(Peto) meta-analysis of above 14 trials: OR=0.57 (95% CI: 0.46,0.71)				
Random Effect (DL) meta-analysis of above 14 trials: OR=0.47 (95%CI: 0.33, 0.68)				
ISIS-4	2216	29011	2103	29039
Fixed Effect(Peto) meta-analysis of above 14 trials: OR=1.0098(95% CI: 0.95,1.07)				
Random Effect (DL) meta-analysis of above 14 trials: OR=0.53 (95%CI:0.36,0.77)				

Note: See the R code for OR calculation in Appendix

3.2 Forest Plot

I produced two plots- one is for log odds ratio, and another one is for odds ratio.



The odds ratio of one indicates that there is no treatment effect. Odds ratio less than one means that there may be some treatment effect. Odds ratio greater than one means the treatment effect may have the opposite effect. Specifically, the 8 trials in the fixed effect model show that the odds ratio is 0.65. This could mean that the odds of being dead in the treatment group is 0.65 times the odds being dead in the control group; Similar interpretation for the 14 trials in the fixed effect model; Their odds ratios are less than 1. This could suggest there is a significant treatment effect. However, in the 15 trials, the odds ratio is greater than 1. It could suggest there is no significant treatment effect after adding ISIS-4 study. I think this is because the ISIS-4 is a large study and it can have a much greater weight in meta-analysis by using Peto method. The forest plot shows the odds ratio for ISIS-4 is 1.06 and the sample size of ISIS-4 is relatively large than all other trials. In Peto method, ISIS-4's study can have greater impact on the overall treatment effect. In contrast, the odd ratio in random effect model does not change dramatically. In details, the odds ratio for the first 8 trials is 0.55 under random effect model. After adding 6 more studies, it decreased a little to 0.47. Adding the large study ISIS-4, the odds ratio only changed from 0.47 to 0.53. These odds ratios in random effect analysis show that there is a significant treatment effect. In conclusion, I prefer random effect model. I think this is because the between study variance is incorporated into the random effect model. So these three tests results did not change much under random effect model.

4. The Bayesian Analysis

Bayesian statistics is an approach to statistics based on a different philosophy from that which underlies significance tests and confidence intervals. It is essentially about updating of evidence. In a Bayesian analysis, initial uncertainty is expressed using a prior distribution^[12]. In this meta-analysis, the prior distribution describes uncertainty regarding the odds ratio. In this Bayesian analysis, I am using Rstan package in R. Here is the model specification:

rc is the number of death in the control group;

pc is the probability of being dead in the control group;

nc is the number of patients in the control group;

rm is the number of death in Magnesium group;

pm is the probability of being dead in the Magnesium group;

nm is the number of patients in the Magnesium group;

We are using the binomial distribution in the stan:

$$P(X = rc) = \binom{nc}{rc} pc^{rc} (1 - pc)^{nc - rc}$$

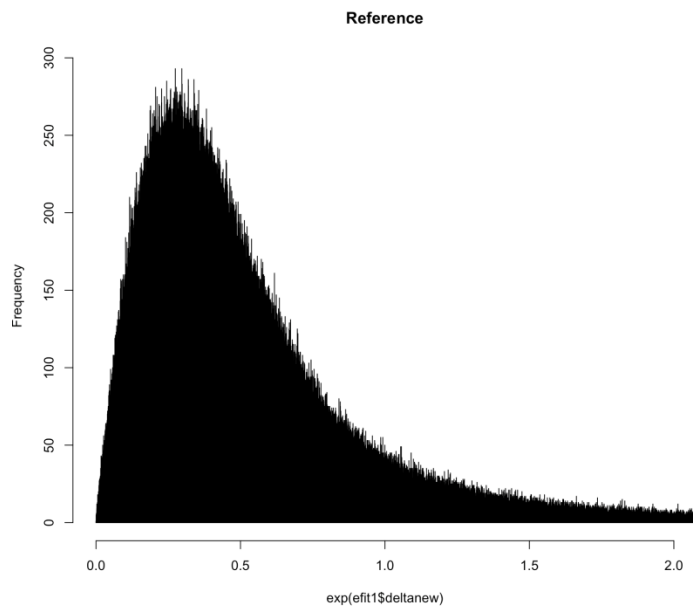
$$P(X = rm_i) = \binom{nm_i}{rm_i} pm_i^{rm_i} (1 - pm_i)^{nm_i - rm_i}$$

The log odds ratio in the Magnesium group, pm_i , is modeled by this:

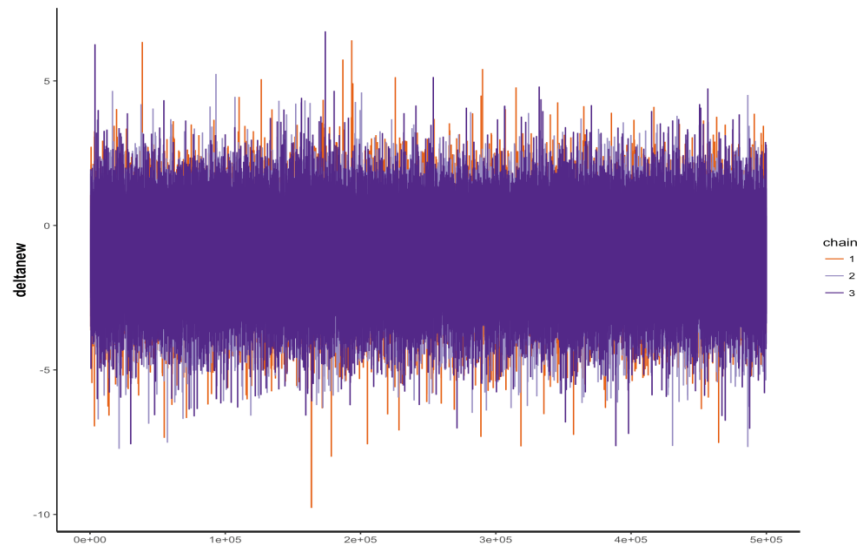
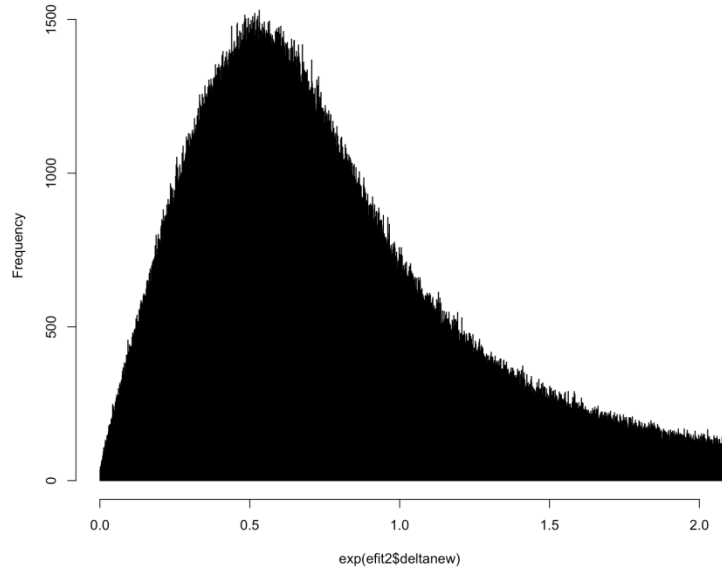
$$\log\left(\frac{pm_i}{1 - pm_i}\right) = \log\left(\frac{pc_i}{1 - pc_i}\right) + \delta_i$$

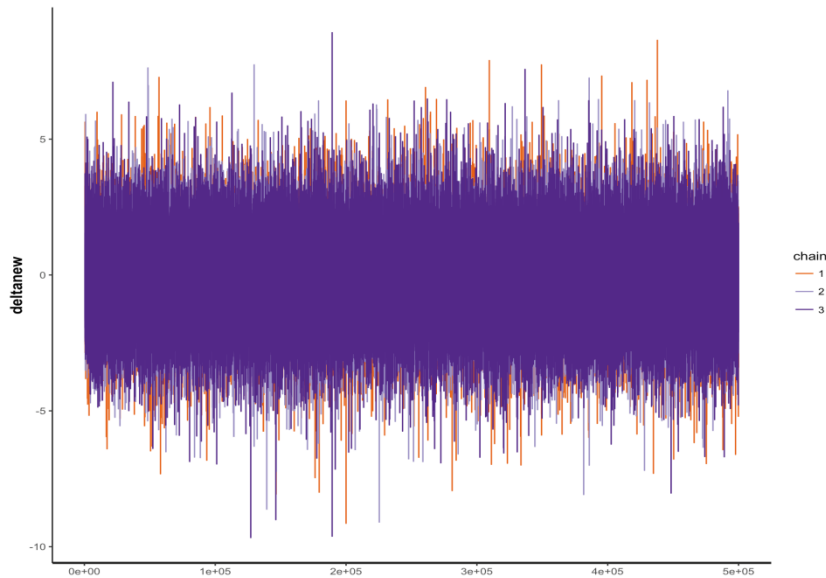
Where $\delta_i \sim N(\mu, \sigma)$, In this model, we specify μ and $\sigma \sim N(0, 100)$. I choose this as our reference prior. However, for our skeptical prior, we let $\mu \sim N(0, 1/\sqrt{32.69})$.

After running the Stan code (attached in the appendix), we can see the estimated mean odds ratio of reference prior is 0.395 ($\exp(-0.93)$, in the Stan model, I got the log odds ratio instead of odds ratio).and skeptical prior is 0.748. Here are the histograms and traceplots for both models:



Skeptical





We can see both reference and skeptical histograms are highly skewed. For reference histogram, most of the odds ratios fall into the 0 to 0.5 range. But for skeptical histogram, most of the odds ratios fall into 0 to 0.8 by eyeballing.

From these two traceplots we can see that these 3 MCMC chains are very consistent. By checking the convergence of these MCMC chains, we can tell both plots have constant mean and variance. For the reference prior distribution, I got the posterior probability of odds ratio less than one is 0.896. The posterior probability of odds ratio less than 0.9 is 0.869. For the skeptical prior distribution, the posterior probability of odds ratio less than one is 0.674. The posterior probability of odds ratio less than 0.9 is 0.62. We can see that the posterior probability for skeptical is less than the reference.

5. Conclusion

In this project, I tried both frequentist analysis and Bayesian analysis. In Frequentist analysis, I have tried with both fixed effect model (Peto) and random effect model (DerSimonian and Laird Method). In Peto method, we can see the obvious limitation of this method after adding the ISIS-4 study, which has much greater sample size and effect size. In contrast, the DerSimonian and Laird is relatively stable even though we have added ISIS-4 study. In the meta-analysis of using all 15 trials, Peto result suggest there is no treatment effect. 15 trials meta-analysis in DerSimonian and Laird method suggest that the odds ratio is smaller than 1, and therefore there is some treatment effect for Magnesium in Myocardial Infarction. This corresponds to Antman's study conclusion that a random effects model gives materially different results from the fixed effect model used by Teo *et al*^[6].

Then I used the Bayesian analysis for these 15 trials. Then I estimated the mean (log) odds ratio for both reference prior and skeptical prior. The mean odds ratio for reference prior is 0.395, and the mean odds ratio for skeptical prior is 0.748. Both results suggest there is some treatment effect in Myocardial Infarction. This means that the odds of being dead in the treatment group is 0.395 times the odds being dead in the control group under the reference prior. Under the skeptical prior, the odds of being dead in the treatment group is 0.748 times the odds being dead in the control group. In addition, we can tell that the posterior probability of odds ratio less than 1 for skeptical prior is much less than the reference prior, which is expected.

Overall, random effect (DL method) has the same results with Bayesian analysis. Their result indicates there is some treatment effect of Magnesium in Myocardial Infarction. I personally would prefer Bayesian analysis. On the one hand, in this project, it has constant result with random effect. On the other hand, I prefer interpreting results using probability. Therefore, I would believe the Bayesian's result of this meta-analysis. This means that magnesium is effective in preventing acute myocardial infarction.

Appendix

```
require(foreign)

#dat<- read.csv("C:/Users/jyao/Downloads/magnesium.csv", header = TRUE)
dat<- read.csv("~/Downloads/magnesium.csv", header = TRUE)
require(metafor)
require(rmeta)

#####peto analysis#####
#peto method for all 15 trials
peto15<-rma.peto(ai=dead1, n1i=tot1,ci=dead0,n2i =tot0, data=dat)
peto15

#peto method for 8 trials
cat<- sample(0,15, replace = TRUE)
dat2<- cbind(dat, cat)
dat2$cat[1:7]<-1
dat2$cat[13]<-1

peto8<-rma.peto(ai=dead1, n1i=tot1,ci=dead0,n2i=tot0, data=dat2, subset = (cat==1))
peto8

#peto method for 14 trails
cat2<- sample(0,15, replace = TRUE)
dat3<- cbind(dat2, cat2)
dat3$cat2[1:14]<-1

peto14<-rma.peto(ai=dead1, n1i=tot1,ci=dead0,n2i=tot0, data=dat3, subset = (cat2==1))
peto14

#alternative way
a<- escalc(measure = "PETO",ai=dead1, n1i=tot1,ci=dead0,n2i=tot0, data=dat, add = 0 )
b<- rma(yi,vi, data=a, method = "FE")

#Forest plot
forest(a$yi,a$vi, slab = dat$trialnam, transf = exp)
forest(a$yi,a$vi, slab = dat$trialnam)

#####DerSimonian Laird analysis#####
#all 15 trials
DL15<- meta.DSL(ntrt = tot1, nctrl = tot0, ptrt = dead1, pctrl = dead0, data = dat)
DL15

#8 trials
DL8<- meta.DSL(ntrt = tot1, nctrl = tot0, ptrt = dead1, pctrl = dead0, data = dat2, subset=(cat==1))
dat99<- dat2[dat2$cat==1,]
DL8<- meta.DSL(ntrt = tot1, nctrl = tot0, ptrt = dead1, pctrl = dead0, data = dat99)
DL8

#14 trials
dat88<- dat3[dat3$cat2==1,]
DL14<-meta.DSL(ntrt = tot1, nctrl = tot0, ptrt = dead1, pctrl = dead0, data = dat88)

#####Bayesian Parts #####

k=15
nc<-dat$tot0
nm<-dat$tot1
rc=dat$dead0
rm=dat$dead1

set.seed(100)
model = "data {
int <lower = 0> k;
```

```

int <lower = 0> nc [k];
int <lower = 0> nm [k];
int <lower = 0> rc [k];
int <lower = 0> rm [k];
}

parameters{
real <lower = 0, upper = 1> pc[k];
vector[k] delta;
real mu;
real <lower = 0> sigma;
real deltanew;
}

transformed parameters {
real <lower = 0, upper = 1> pm[k];
for (i in 1:k) {
pm[i] = exp(log(pc[i]/(1-pc[i])) + delta[i])/(1+exp(log(pc[i]/(1-pc[i])) + delta[i]));
}
}

model {
# model
#for (i in 1:k) {
rc ~ binomial(nc, pc);
rm ~ binomial(nm, pm);
delta ~ normal(mu, sigma);
#}
deltanew ~ normal(mu, sigma);

# priors
#for (i in 1:k) {
pc ~ uniform(0,1);
#}

mu ~ normal(0, 100);
sigma ~ uniform(0, 100);

}"

fit1 = stan(model_code = model, data = c("k", "nc", "nm", "rc", "rm"), pars = "deltanew", iter = 500000,
chains = 3, warmup = 500)
print(fit1)
traceplot(fit1, pars= "deltanew")

#odd ratio =0.3945

#####SKEPTICAL#####

model = "data {
int <lower = 0> k;
int <lower = 0> nc [k];
int <lower = 0> nm [k];
int <lower = 0> rc [k];
int <lower = 0> rm [k];
}

parameters{
real <lower = 0, upper = 1> pc[k];
vector[k] delta;
real mu;
real <lower = 0> sigma;
real deltanew;
}

transformed parameters {

```

```

real <lower = 0, upper = 1> pm[k];
for (i in 1:k) {
pm[i] = exp(log(pc[i]/(1-pc[i])) + delta[i])/(1+exp(log(pc[i]/(1-pc[i])) + delta[i]));
}
}

model {
# model
#for (i in 1:k) {
rc ~ binomial(nc, pc);
rm ~ binomial(nm, pm);
delta ~ normal(mu, sigma);
#}
deltanew ~ normal(mu, sigma);

# priors
#for (i in 1:k) {
pc ~ uniform(0,1);
#}

mu ~ normal(0, 0.1749);
sigma ~ uniform(0, 100);

}"

fit2 = stan(model_code = model, data = c("k", "nc", "nm", "rc", "rm"), pars = "deltanew", iter = 500000,
chains = 3, warmup = 500)
print(fit2)
#odd ratio =0.74
traceplot(fit2, pars= "deltanew")

efit1 <- extract (fit1, permuted=TRUE)
hist(exp(efit1$deltanew), xlim =c(0,2), breaks = 10000000, main="Reference")

efit2 <- extract (fit2, permuted=TRUE)
hist(exp(efit2$deltanew), xlim = c(0,2), breaks = 10000000, main="Skeptical")

```

Reference

1. Russo, Mark W. "How to Review a Meta-analysis." *Gastroenterology & Hepatology*. Millennium Medical Publishing, Aug. 2007. Web. 11 May 2017.
2. "Study Design 101." *Study Design 101 - Meta-Analysis*. N.p., n.d. Web. 11 May 2017.
3. Dechartres, Agnes, Ludovic Trinquart, Isabelle Boutron, and Philippe Ravaud. "Influence of trial sample size on treatment effect estimates: meta-epidemiological study." *The BMJ*. BMJ Publishing Group Ltd., 2013. Web. 11 May 2017.
4. Orloff, Jeremy, and Jonathan Bloom. "Comparison of Frequentist and Bayesian Inference Class 20, 18.05." *MITOPENCOURSEWARE*. N.p., n.d. Web. <https://ocw.mit.edu/courses/mathematics/18-05-introduction-to-probability-and-statistics-spring-2014/readings/MIT18_05S14_Reading20.pdf>.
5. Parker, Mary. "Foundations of Statistics- Frequentist and Bayesian." *Austin Community College*. N.p., n.d. Web. <<http://www.austincc.edu/mparker/stat/nov04/>>.
6. Higgins, J. P., and D. J. Spiegelhalter. "Being sceptical about meta-analyses: a Bayesian perspective on magnesium trials in myocardial infarction." *International journal of epidemiology*. U.S. National Library of Medicine, Feb. 2002. Web. 11 May 2017.
7. "Meta-analysis." *Wikipedia*. Wikimedia Foundation, 09 May 2017. Web. 11 May 2017.
8. Borenstein, Michael, Larry Hedges, and Hannah Rothstein. "Meta-Analysis: Fixed Effect vs. Random Effects." *Comprehensive Meta-Analysis*. N.p., 2007. Web. 11 May 2017. <https://www.meta-analysis.com/downloads/M-a_f_e_v_r_e_sv.pdf>.
9. "Peto Odds Ratio Meta-analysis." *Peto Odds Ratio Meta-analysis - StatsDirect*. N.p., n.d. Web. 11 May 2017.
10. Brockhaus, A. C., R. Bender, and G. Skipka. "The Peto odds ratio viewed as a new effect measure." *Statistics in medicine*. U.S. National Library of Medicine, 10 Dec. 2014. Web. 11 May 2017.
11. Dan Jackson, Jack Bowden, Rose Dawn Bake. "How does the DerSimonian and Laird procedure for random effects meta-analysis compare with its more efficient but harder to compute counterparts?" *ResearchGate*. N.p., 01 Apr. 2010. Web. 11 May 2017.
12. *16.8.1 Bayesian methods*. N.p., n.d. Web. 12 May 2017. <http://handbook.cochrane.org/chapter_16/16_8_1_bayesian_methods.htm>.