Chuan Ya Hsu
August 1, 2017
APSTA-GE 2122
Applied Statistical Modeling and Inference
Instructors: Ying Lu
Daphna Harel

Final Project

Introduction

Meta-Analysis

The belief of meta-analysis is that a common truth exists behind all conceptually similar scientific studies. Nevertheless, the truth has been measured with certain error between each study. Hence, the purpose of meta-analysis is to produce a pooled estimate, which approaches to the unknown common truth by using statistical ways, and based on how the error is perceived.

To obtain the information from studies in order to produce the estimation of unknown common truth, meta-analysis is capable of comparing the results from various studies, and clarifying which contexts are the interested targets. Meta-analysis, in other words, can be thought as a kind of literature review.

The advantage of this approach is that the aggregation information it has been able to cause a higher statistical power, and the point estimate it generates is more robust than those yields by any individual study. [1]

Factors that Cause Problem in Meta-Analysis

In a meta-analysis, there are some factors which researcher must concern, for they might cause serious bias. Some of the most crucial factors are the number of trials (studies) which been included, the quality of the study, the size of each study, and the file drawer problem.

Some researchers have debated that the fact that a meta-analysis cannot control the existence of sources of bias. Methodologically believes that the condition of "garbage in, garbage out" may happen if the analysis contains studies without selecting. Other meta-analysts include weak studies with adding a study-level predictor variable which reports the quality of studies to examine the influence of studies on the effect size. Others, however, grab as widely types of studies as they can, and state that remaining information about the variance in the studies themselves is a better way.

For the size of the studies, the smaller studies generally have larger standard errors, and

might be less accurate. If considering the file drawer problem (i.e. publication bias) combines with the study size, more issue will need to be reviewed. For instance, if an analysis has no publication bias, there were no relationship between standard error and effect size. In the scatter plot of standard error versus the effect size, smaller studies which found effects in only one direction indicates that it is likely only the negative or positive studies had been published. [1]

The Differences Between Bayesian and Frequentist Views of Statistics

Before discussing the various methods which we will display in the later parts, the distinct fundamental concepts of frequentist and Bayesians views of statistics are worth it to be reviewed.

For the view of data, frequentist believes that data are repeatable random samples, which the frequency is exist and the studies are repeatable. Nonetheless, it treats the repeatability as the most vital thing, whatever we pay for it. Bayesian, on the other hand, considers data as the observation of the realized sample (i.e. the studies are fixed). In other words, Bayesian believes that data are fixed.

In terms of the parameters, frequentist trusts that the parameters are fixed, which indicates that the parameters will remain constant during the repeated process, and even under all circumstances. On the contrary, Bayesian views the world probabilistically, which leads to the result that it considers parameters are unknown and described probabilistically.

Last but not the least, there is no information prior to the model specification for frequentist; while Bayesian has prior information, and regards them as the most important part in the process. Furthermore, the assumption that data is from a controlled experiment is made in frequentist whereas Bayesian is careful of prescribing assumptions. In general, frequentist tends to believe that the information and methods we interest and use are objective yet Bayesian sees every statistical model ever created in the history are subjective. [2] [3]

Meta-Analysis and Bayesian Paradigm

As the discussion in previous part, several factors will lead the meta-analysis of bias results. For instance, the previous analysis shown in Higgins and Spiegelhalter's paper interpreted a large treatment effect of magnesium. Nevertheless, the treatment effects, in our suspicion, should be more moderate. The idea of Bayesian, which allows researchers to set a prior, can achieve our assumption well in this case, for we are able to set a prior which reflects our belief that treatment effects are unlikely to be so large a value. Hence, the treatment effects generate by Bayesian model are consistent with our belief and more reasonable to interpret.

For a long time, both physiologists' studies in animals and humans, and epidemiologists' studies had recommended that intravenous magnesium was crucial in patients with acute myocardial infraction. However, after further analysis under mega-trial, Yusuf, who is the main author of the optimistic Circulation editorial, claimed that the results of previous studies might "too good to be true," for it seemed that the reducing morality effect for routine magnesium using should tend to be moderate rather than the extraordinary large reduction showed in former studies.

Instruction of Models

The Philosophical Difference Between Fixed Effects and Random Effects Meta-Analysis

In the design of fixed effects meta-analysis, there is one true effect size which underlies all studies. The method considers that all differences between studies are simply random errors. Since the confidence interval is based on each study, which only affected by the sample size, with a constant central point estimation, the confidence interval is relatively narrow.

Different from fixed effects meta-analysis, random effects meta-analysis is mainly built on the basis of heterogeneity. The method follows the concepts that study effect is from a wide distribution of study effects, and it is impossible that studies having a common effect size for all studies are naturally various. The above conditions lead random effects meta-analysis to have a wider confidence interval, and concern not only within but also between study variance. [4]

Study Size and Overall Treatment Effect

It is regular that a meta-analysis involves many small studies within it. A small study does not capable of detecting a fair intervention effect. It inclines to inform greater intervention effect, compare with the larger study. Individuals, small studies show more extreme treatment effects than large studies. However, when discussing the pooled treatment effect, the issue of weighting each study must be concerned especially for fixed effects models. In fixed effects meta-analysis, the weight of each study depends on their size. A larger study will have a greater weight than the small study. As the result, the sizes of studies decide how "important" the studies are in the pooled treatment effect produced by fixed effects meta-analysis. [5] [6]

Overview of the Dataset

The trials in the dataset have basically followed the order of providing time (year). Most

of them have under 150 patients in the treatment group and a similar number in the control group, which the sizes are rather small comparing with the especially large two trials. LIMIT-2 has 1,159 patients in the treatment group, and 1,157 patients in the control group. The largest trial ISIS-4 has 29,011 patients in the treatment group, and 29,039 patients in the control group. It is no doubt ISIS-4 will influence the fixed effects method because of the fact that the weight of ISIS-4 will larger than others.

The Peto Method

Peto's method can only be used to pool odds ratios. Its approximation of the log odds ratio works well when the odds ratios are close to one, events are not particularly common, and the treatment and control groups have a similar size. Nevertheless, Peto's method can cause bias in other cases, especially when the sizes of treatment and control groups are substantially different.

The estimation way of odds ratio according to Peto's method is as follows:

$$O_i = r_i^M,$$

$$E_i = \frac{\left(r_i^M + r_i^C\right)\left(r_i^M + n_i^M - r_i^M\right)}{n_i^M + n_i^C} = \frac{\left(r_i^M + r_i^C\right)n_i^M}{n_i^M + n_i^C},$$

$$V_i = \frac{\left(r_i^M + r_i^C\right)\left(n_i^M + n_i^C - r_i^M - r_i^C\right)n_i^M n_i^C}{\left(n_i^M + n_i^C\right)^2\left(n_i^M + n_i^C - 1\right)},$$

$$\hat{\psi}_{pool} = \exp\left(\frac{\sum_{i=1}^{k}(O_i - E_i)}{\sum_{i=1}^{k} V_i}\right),$$

$$CI_{pool} = \exp\left(\frac{\sum_{i=1}^{k}(O_i - E_i) \pm Z_{\alpha/2}\sqrt{\sum_{i=1}^{k} V_i}}{\sum_{i=1}^{k} V_i}\right),$$

where O_i is the observation of death number in treatment group for the i[th] trial;
E_i is the expected death number in treatment group for the i[th] trial;
V_i is both weighting factor and variance for the difference between observed and expected, i.e. $O_i - E_i$;
k is the number of trials;
$\hat{\psi}_{pool}$ is the pooled Peto odds ratio for all trials;
and CI_{pool} is the confidence interval of $\hat{\psi}_{pool}$. [5] [6]

Chuan Ya Hsu
August 1, 2017

The DerSimonian and Laird (D-L) method

The random effects models are under the assumption that the treatment effect from the i[th] trial, Y_i, is from the distribution $Y_i|\mu_i \sim N(\mu_i, \sigma_i^2)$, where μ_i is the real treatment effect of i[th] trial, and σ_i^2 is the corresponding within variance. The models further assume that $\mu_i \sim N(\mu, \tau^2)$, where μ and τ^2 represent the overall treatment effect and between-study variance respectively.

DerSimonian and Laird (D-L) method is a well-known simple way to evaluate the between-study variance of random effects meta-analysis, which uses the Q statistic

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

where $w_i = \sigma_i^2$, $\bar{y} = \frac{\sum_{i=1}^{k} w_i y_i}{\sum_{i=1}^{k} w_i}$, and k indicates the number of trials.

By the definition of random effects model, the expected value of Q statistic can be expressed as

$$\mathrm{E(Q)} = (\mathrm{k} - 1) + \left(S_1 - \frac{S_2}{S_1}\right)\tau^2,$$

where $S_r = \sum_{i=1}^{k} w_i^r$, which offers the DerSimonian and Laird estimate

$$\hat{\tau}_{DL}^2 = \max\left(0, \frac{Q - (k-1)}{S_1 - \frac{S_2}{S_1}}\right).$$

The corresponding estimation of treatment effect is

$$\hat{\mu}_{DL} = \frac{\sum_{i=1}^{k} \frac{y_i}{\sigma_i^2 + \hat{\tau}_{DL}^2}}{\sum_{i=1}^{k} \frac{1}{\sigma_i^2 + \hat{\tau}_{DL}^2}},$$

and the confidence interval of it follows the distribution $\hat{\mu}_{DL} \sim N(\mu, \frac{1}{\sum_{i=1}^{k} w_i^*})$, which $w_i^* = \frac{1}{\sigma_i^2 + \hat{\tau}_{DL}^2}$. Hence the $100(1-\alpha)\%$ confidence interval for μ is

$$\hat{\mu}_{DL} \pm Z_{\alpha/2} \frac{1}{\sqrt{\sum_{i=1}^{k} w_i^*}}. \quad [7]$$

The Frequentist Analysis

Table 1: Reproduced table from table 2 in Higgins and Spiegelhalter paper.

	Magnesium group		Control group	
Trial	Deaths rm	Patients nm	Deaths rc	Patients nc
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
Schechter	1	59	9	56
Ceremuzynski	1	25	3	23
LIMIT-2	90	1159	118	1157
Fixed effect (Peto) meta-analysis of above eight trials: OR=0.65 (95% CI: 0.51, 0.82);				
Random effects (D-L) meta-analysis of above eight 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
Schechter 2	4	107	17	108
Fixed effect (Peto) meta-analysis of above 14 trials: OR=0.57 (95% CI: 0.46, 0.71);				
Random effects (D-L) meta-analysis of above 14 trials: OR=0.47 (95% CI: 0.32, 0.68);				
ISIS-4	2216	29011	2103	29039
Fixed effect (Peto) meta-analysis of above 15 trials: OR=1.01 (95% CI: 0.95, 1.07);				
Random effects (D-L) meta-analysis of above 15 trials: OR=0.53 (95% CI: 0.36, 0.77);				

The odds for a group (treatment or control) is defined as the death rate of the patients divided by the survive rate. Odds ratio, consists with its name, is the ratio of two odds, which implies that the ratio of the odds of treatment group to control group. If the odds ratio is less than one, there is a decreased likelihood of death rate in the treatment group. On the other hand, if the odds ratio is equal to one, it means that the likelihood of death rate in the treatment group is no different compared with the control group. In addition, if the odds ratio is greater than one, it indicates that the likelihood of death rate increases from the control group to treatment group. [7]

In the reproduced output table, we notice that the pooled odds ratios produced by fixed effects meta-analysis are 0.65, 0.57, and 1.01 respectively. This result demonstrate that magnesium could effectively reduce the death rate if we do not include the last large trial. If ISIS-4 is involved in the analysis, the reduction of death rate affected by magnesium become invisible. Moreover, magnesium might even increase the death rate under the analysis result.

Though the conclusions based on fixed effects change between models, the conclusions judging by random effects do not change after including more trials in the model. Both three odds ratios display that magnesium is able to decrease the death rate, and the odds ratios produced by random effects seem to be more consistent compare with the fixed effects model. Additionally, the assumption which random effects makes, that every student or trials are different in nature, meets the real condition of our dataset better because we cannot ignore the fact that how various the size of the trials are. As the result, the results from random

effects models appear to be more reliable.

In another aspect, the influence for fixed effects meta-analysis from trials' sizes can be clearly seen on the forest plots. The pooled log odds ratio of fist eight trials, as Figure 1 shown, is -0.44. The value is quite close to the log odds ratio of trial LIMIT-2, the largest trial in the first eight. The pooled log odds ratio of the analysis contains first fourteen trials is still near the log odds ratio of LIMIT-2, which can be seen in Figure 2, though being pulled away a little by other trials. In Figure 3, the pooled log odds ratio presents an enormous change from -0.56 to 0.01 when the analysis involve the large trial ISIS-4. Interestingly, the log odds ratio of ISIS-4 is very close to zero, which seems affected a great deal of the whole analysis.

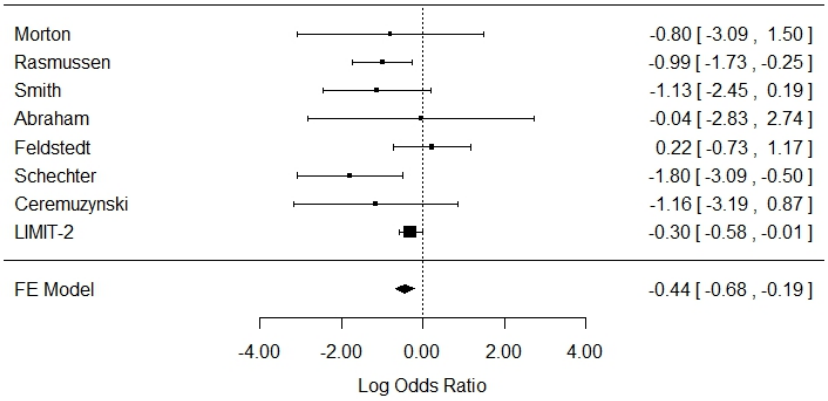

Figure 1: Fixed effect (Peto) meta-analysis of the first 8 trials

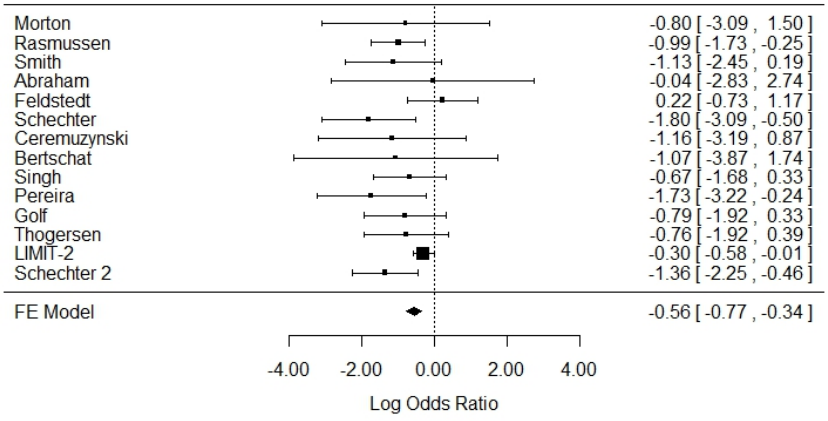

Figure 2: Fixed effect (Peto) meta-analysis of the first 14 trials

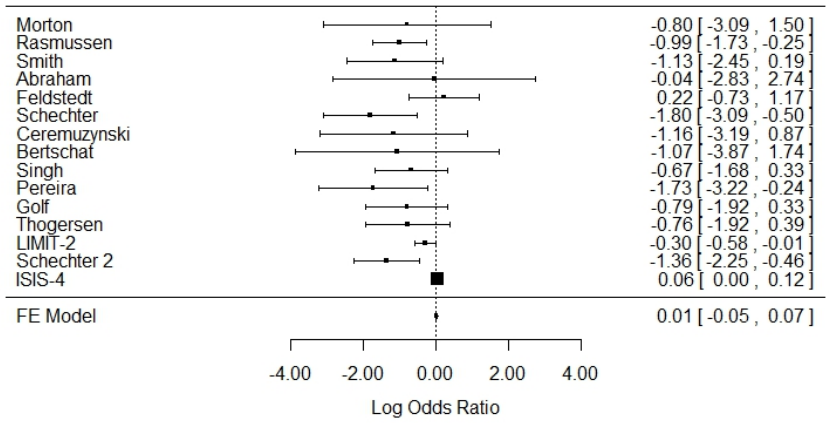

Figure 3: Fixed effect (Peto) meta-analysis of the 15 trials

The Bayesian Analysis

For the Bayesian random effects meta-analysis models under the reference prior, we assume that

$$\mu \sim N(0, 100),$$

$$\sigma^2 \sim Uniform(0,100), \text{and}$$

$$p_i^C \sim Uniform(0,1),$$

where μ is the expected value of estimated treatment effects, σ^2 is the variance of estimated treatment effects, and p_i^C is the expected probability of r_i^C, i.e.

$$r_i^C \sim Binomial\left(n_i^C, p_i^C\right),$$

$$estimated \log(odds\ ratio) = \delta_{new} \sim N(\mu, \sigma^2).$$

By using 3 MCMC chains, with 50,000 iterations on each chain[1], we get the estimated log odds ratio. However, without checking whether the chain has converged to its stationary distribution, we cannot sure the if the chain is mixing well. [11] As the result, Figure 4, which is the trace plot of δ_{new}, provide a visualized examine of the chain.

[1] The reason why I do not use at least 500,000 iterations on each chain was because my computer could not handle such enormous process.

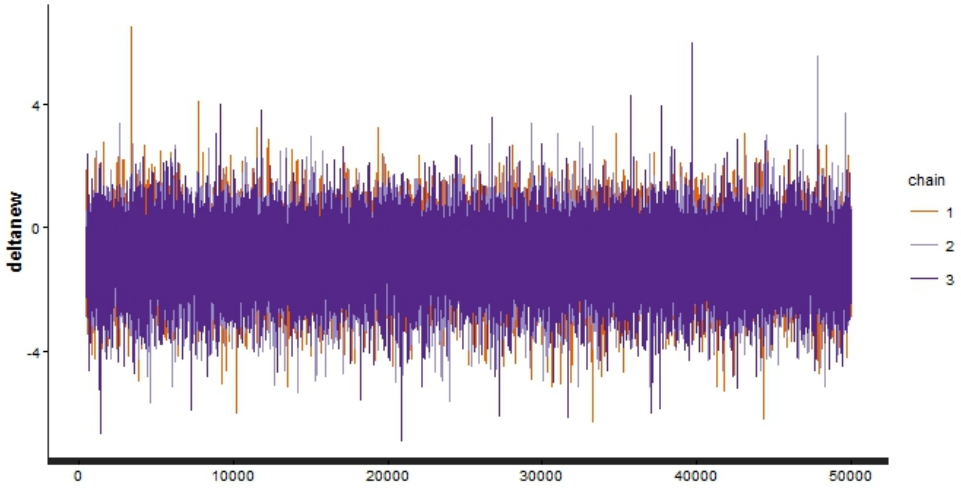

Figure 4: Trace plot of δ_{new} produced by reference prior distribution

The Figure 4 displayed a "perfect" trace plot, which make us sure that the chain has converged. We continue observing the summary statistics of estimated odds ratio, and obtain the following information:

Mean	Minimum	1st Quantile	Median	3rd Quantile	Maximum
0.57	0.001	0.25	-0.90	0.65	694.10

Table 2: Summary statistics of estimated odds ratio based on reference prior.

On the other hand, we only change the prior setting distribution of μ to

$$\mu \sim N(0, 0.1749),$$

where $0.1749 = \frac{1}{32.69} = \frac{1}{\tau^2}$, and the τ^2 had been set by Higgins and Spiegelhalter in their paper, in the skeptical prior Bayesian model.

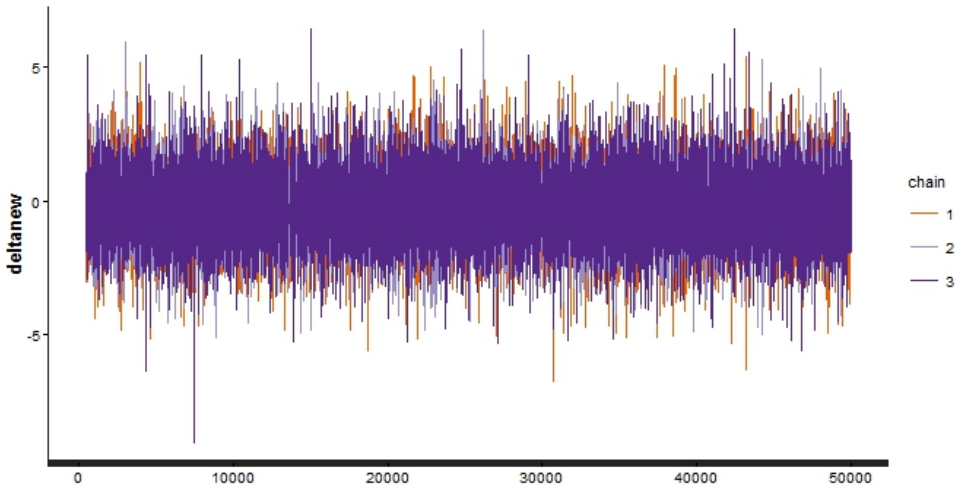

Figure 5: Trace plot of δ_{new} produced by skeptical prior distribution

As the above process, we check whether the chain converged or no. The result is shown in Figure 5. Based on the figure, the chain has already converged to a stationary distribution. Hence, we can further generate the basic statistics of estimated odds ratio, which represent in Table 3.

Mean	Minimum	1st Quantile	Median	3rd Quantile	Maximum
0.57	0.001	0.25	-0.90	0.65	694.10

Table 3: Summary statistics of estimated odds ratio based on skeptical prior.

Furthermore, the histograms of the posterior distribution for the overall meta-analytic estimate of the odds ratio for both Bayesian models can also illustrate the summary statistics. Since the log odds ratios follow the prior setting, which normally distributed with the μ and σ^2, the distributions of odds ratios are skewed to right due to the exponential transform.

In Figure 6, most of the estimates are located between 0 and 1, which suggest that the treatment effects appear to be significant under the reference prior distribution. The histogram of the posterior distribution for estimated odds ratio under skeptical prior distribution is shown in Figure 7. Although most of them still spread between 0 and 1, the model under skeptical prior distribution display a higher probability that there is no significantly treatment effects.

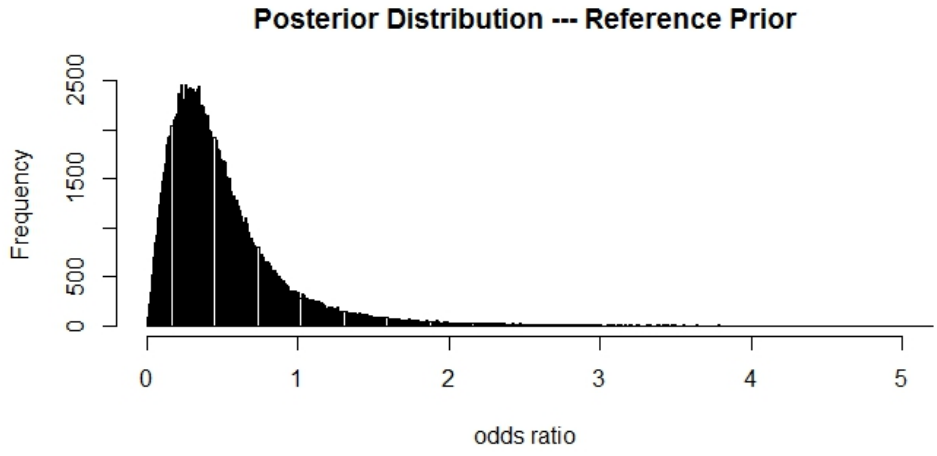

Figure 6: Posterior distribution for the overall meta-analytic estimate of the odds ratio under reference prior.

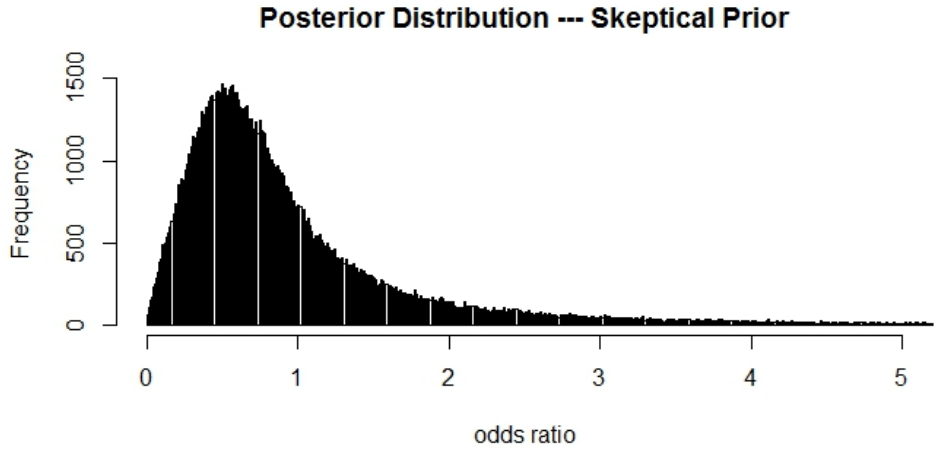

Figure 5: Posterior distribution for the overall meta-analytic estimate of the odds ratio under skeptical prior.

Table 4 demonstrates that the Bayesian model under reference prior distribution finds the evidence quite reliable based on the clinical improvement, which there is a 87% probability that magnesium can reduce at least 10 % of the mortality rate. Model under skeptical prior distribution, however, claims that the probability of magnesium decrease at least 10% death rate is only about 62 %.

	Magnesium superior P(odds ratio < 1)	**Magnesium clinically superior P(odds ratio < 0.9)**
Reference Prior	0.8967	0.8710
Skeptical Prior	0.6744	0.6212

Table 4: Posterior probabilities of statistical and clinical superiority of magnesium, given two prior distributions

Conclusion

In this project, we come through the frequentist fixed effects meta-analysis, frequentist random effects meta-analysis, Bayesian random effects meta-analysis under reference prior distribution, and Bayesian random effects meta-analysis under skeptical prior distribution. The results of them are somehow distinguishing owing to the various assumptions they have.

For frequentist views of statistics, we observe that the result of fixed effects meta-analysis is extremely sensitive to the sizes of the trials, which makes the result analysis unstable. When the analysis excludes the large trial ISIS-4, the treatment effect it produces is similar with a random effects model. However, the treatment effect changes in a great deal after the ISIS-4 been included in the analysis. Random effects model, yet does not

significantly influenced by the study size. Nonetheless, both two models perform a rather high treatment benefit. These results had been argued as “too good” for the reality because they both present that magnesium has the ability to reduce about half of the mortality rate.

The two Bayesian models demonstrate the results more consist with our assumption, for both two prior settings are follow our expectation. However, the result of model under reference prior distribution may still too idealise for it is hard for a treatment to perform so steady under any circumstances. There must have a wide variance of mortality reduction rate between each case. Hence, the most appropriate method for this analysis is the Bayesian model under skeptical prior distribution because the suspicion of the treatment effect should be more moderate has already been concerned with the setting of this method.

Many methods showed that magnesium can effectively lower the death rate of acute myocardial infarction, the possible probability they generated of high reduction is unfortunately unrealistic. In this case, the probability of magnesium decreases at least 10% death rate is around 62 %, which estimated under skeptical prior distribution, sounds more reliable compare with other methods. Frankly speaking, we cannot deny magnesium must prevent acute myocardial infarction to some extent. However, in what degree magnesium is capable of decreasing the death rate still needs further studies.

Reference

[1] "Meta-Analysis," [Online]. Available: https://en.wikipedia.org/wiki/Meta-analysis.

[2] Berger and Berry, “Statistical Analysis and the Illusion of Objectivity,” *American Scientist,* 1998.

[3] G. Casella, “Bayesians and Frequentists,” ACCP 37th Annual Meeting, Philadelphia, PA.

[4] T. Shaneyfelt, *Fixed Effects and Random Effects Models.* [Video]. 2013.

[5] S. M. B. J. P. T. H. Rebecca M. Turner, "The Impact of Study Size on Meta-analysis: Examination of Underpowered Studies in Cochrane Reviews," 2013.

[6] M. L. H. a. H. R. Borenstein, "Meta-analysis: Fixed effect vs. random effects," Meta-Analysis.com, 2007.

[7] "Peto Odds Ratio Meta-analysis," StatsDirect Limited, [Online]. Available: http://www.statsdirect.com/help/default.htm#meta_analysis/peto.htm.

[8] Julian PT Higgins and Sally Green, "Cochrane Handbook for Systematic Reviews of Interventions 9.4.4.2 Peto odds ratio method," [Online]. Available: http://handbook.cochrane.org/chapter_9/9_4_4_2_peto_odds_ratio_method.htm.

[9] Dan Jackson, Jack Bowden, Rose Baker, "How does the DerSimonian and Laird procedure for random effects meta-analysis compare with its more efficient but harder to

compute counterparts?," *Journal of Statistical Planning and Inference,* pp. 961-970, 2009.

[10] A. K. Akobeng, "Understanding Systematic Reviews and Meta-Analysis," *Arch Dis Child,* vol. 90, Issue 8, 2005.

[11] "Introduction to Bayesian Analysis Procedures: Assessing Markov Chain Convergence :: SAS/STAT(R) 9.2 User's Guide, Second Edition," SAS/STAT(R), [Online]. Available: https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_introbayes_sect008.htm.