On Bayes and Frequentist Meta-Analyses

Hal Li[1], William Wang[1]

[1]Merck Research Laboratories, 351 N Sumneytown Pike, North Wales, PA 19454

Abstract

The motivation is to examine the clinical applications of Bayesian theory in meta-analysis and to develop methods of applying this theory to safety data. We developed and applied a hybrid approach that combines Bayesian hierarchical modeling with the frequentist approach to meta-analysis. The question with regards to selecting a likelihood function will be presented. The analysis is conducted using the example of a meta-analysis regarding risk management safety data. The hybrid analysis is derived by verifying assumptions through a frequentist approach. The implications of this combined approach are discussed, including using Q heterogeneity statistics and diagnostic plots to determine exchangeability. This analysis will also address the process of determining the proper Bayesian prior and likelihood distributions. Finally, Bayesian methods are illustrated using graphical methods to examine extreme probabilities in safety data. The proper Bayesian distribution is determined through these safety graphs.

Key Words: Safety Assessment, Cumulative Safety Data Meta-analysis, Bayesian hierarchical modeling, Safety Graphs

1. Motivating Example

Developing a potential application of Bayes and frequentist meta-analyses methods was the objective of this study, which sought to analyze a Bayesian meta-analysis using hierarchical modeling based on different Likelihood Functions (i.e. different distributions). In addition, Bayesian meta-analysis was compared to traditional Peto odds ratio meta-analysis.

The motivating example of this meta-analysis was to evaluate the prevalence of oropharyngeal adverse events (Candidiasis, Pharyngitis, Dysphonia, Cough) induced by different inhaled corticosteroids (ICS). These effects have been studied less extensively than those that occur systemically and thus provides an open area of investigation.

1.1 Search Strategy

A computerized search in the MEDLINE (January 1966 to June 2016) and EMBASE (January 1974 to June 2016) databases was conducted using appropriately indexed MedDRA terms. These terms included the following: candidiasis, dysphonia, hoarseness, pharyngitis, thrush, throat irritation, voice alteration / dysfunction, distorted voice, laryngeal/pharyngeal pain, oral fungal infection, cough, oropharyngeal / esophageal adverse event, dose, local safety, incidence, prevalence, epidemiology, spacer, aerosol, asthma, and inhaled corticosteroids..

1.2 Search Strategy

Only randomized, placebo-controlled studies, with an emphasis on oral ICSs (single entity or combination therapy) for the treatment of persistent asthma of all severities, were eligible for inclusion (chronic obstructive pulmonary disease was excluded).

The included studies were efficacy and safety studies not specifically designed to report local safety. Study populations were limited to adults and adolescent cohorts only. Furthermore, only studies that reported appropriate information on patient demographics and study design were included.

A total of 50 studies was selected based on the criteria detailed above.

2. Bayesian and Meta-Analysis

2.1 **An Overview of the Bayesian Approach**

Supposing that we are interested in estimating θ from a data set: $X=\{x_1 ,\ldots,x_n\}$. Bayes theorem provides a solution by using a simple well known rule about conditional probabilities:

$$P(\theta_i \mid X) = \frac{P(X,\theta_i)}{P(X)} = \frac{P(X \mid \theta_i)P(\theta_i)}{\sum_j P(X \mid \theta_j)P(\theta_j)} \qquad [1]$$

Overall, Bayesian inference is based on the posterior distribution of the parameter P(θ|X). However, in order to derive the posterior distribution, we need to specify the prior distribution, P(θ) – the distribution of θ, we also need to determine the likelihood function P(X|θ) from the data observed. From [1], one can see that the P(θ|X) is proportional to (i.e. has the same shape as) the product of the likelihood function and the prior distribution of the data:

$$P(\theta|X) \sim P(X|\theta)P(\theta) \qquad [2]$$

Having derived the posterior distribution, P(θ|X), in Bayesian analysis all further inferences about θ will be derived from that distribution. This includes calculations of location parameters including the posterior mean, median, or mode, or percentiles, among other parameters and etc.

2.2 **MCMC Estimation Method**

In theory, Bayesian methods are straightforward. The posterior distribution contains everything needed to carry out inference. In practice, the posterior distribution can be difficult to estimate precisely. One popular and very general simulation method is a Markov Chain Monte Carlo (MCMC). MCMC methods sample successively from a target distribution, with each sample depending on the previous one, hence the notion of the Markov chain. Monte Carlo integration computes an expectation by averaging the Markov chain samples

$$\int_S g(\theta)p(\theta)d\theta \cong \frac{1}{n}\sum_{t=1}^{n} g(\theta^t)$$

where g(.) is a function of interest and θ^t are samples from p(θ) on its support S.

The SAS proc MCMC procedure was applied for this meta analysis. The MCMC procedure uses a special case of the MCMC method, the random walk Metropolis algorithm (Metropolis et al. 1953; Hastings 1970) {1}, {2} to generate a sequence of draws from the joint posterior distribution of the parameters. This procedure is capable of constructing an optimal proposal distribution in the random walk Metropolis algorithm and this procedure can be used to generate samples from an arbitrary density. Once samples are obtained, one can carry out additional statistical inference as desired.

3. Bayesian Hierarchical Modeling for Meta-Analysis

Meta-analysis is an important technique that combines information from different studies. When there is no prior information for thinking any particular study is different from another, Bayesian meta-analysis can be treated as a hierarchical model. This assumption, known as exchangeability, then allows for a Bayesian random-effects model, which assumes that there is no such prior information and therefore treats all studies equally.

3.1 Normal Approximation to the Binomial Likelihood

Normal approximation to binomial likelihood is a classical method that is commonly used in meta-analysis. Assume that there are total N_trt_j and N_ctrl_j patients in the jth study. Assume also that the number of oral candidiasis events observed in the jth study active treatment arm and control arm as $event\_trt_j$ and $event\_ctrl_j$ respectively. Given, these assumptions, let Y_j be the odds ratio of oral candidiasis for the jth study:

$$Y_j = \frac{envt\_trt_j(N\_ctrl_j - even\_ctrl_j)}{envt\_ctrl_j(N\_trt_j - even\_trt_j)} \quad [3]$$

It is possible to estimate the treatment effect θ_j, j=1,…,n, through the means of an approximate normal distribution,

$\log(Y_j) \sim \text{normal}(\theta_j, s_j^2)$

where θ_j is the study-specific effect and s_j^2 is the known variance of $\log(Y_j)$. For s_j^2, If all studies have large sample sizes (more than 30 persons in each group in nearly all of the studies), it is possible to use the approximate sampling variance of θ_j:

$$s_j^2 = \frac{1}{even\_trt_j} + \frac{1}{N\_trt_j - even\_trt_j} + \frac{1}{envt\_ctrl_j} + \frac{1}{N\_ctrl_j - even\_ctrl_j} \quad [4]$$

If the odds ratios are exchangeable between the studies, then the treatment effect in each trial can be considered to be a random quantity drawn from some population distribution. In a Bayesian framework, this means that it is possible to place a common prior distribution on the exchangeable random-effects parameters θ_j:j=1,…,n.

$\theta_j \sim \text{normal}(\mu, \tau^2)$

where μ is the population average of the treatment effect across all studies and τ^2 is the between-study variation. In this meta analysis, the following non informative priors are placed on the hyper parameters μ, and τ^2:

$\mu \sim \text{normal}(0, 3^2)$
$\tau^2 \sim \text{igamma}(\text{shape} = 0.01, \text{scale} = 0.01)$

3.2 Bayesian Hierarchical Modeling Use Binomial Likelihood

Emulating the binomial model as outlined in Spiegelhalter, Abrams, and Myles (2004) {3}, one can use the following model, where the treatment and control groups have their own binomial likelihood functions, with oral candidiasis probabilities p_j and q_j in the treatment group and control group respectively:

$\text{event\_trt}_j \sim \text{binomial}(\text{N\_trt}_j, p_j)$
$\text{event\_ctrl}_j \sim \text{binomial}(\text{N\_ctrl}_j, q_j)$

Let the log odds for the control group be ϕ_j and let the log odds for the treatment group be $\theta_j + \phi_j$ as follows:

$\phi_j = \log(q_j/(1- q_j))$
$\theta_j = \log(p_j/(1- p_j))$

Then θ_j is the log odds ratio:

$$\theta_j = \log\left(\frac{p_j(1 - q_j)}{q_j(1 - p_j)}\right) \quad [5]$$

3.3 Bayesian Hierarchical Modeling Use Poisson Likelihood

Similar to the binomial likelihood, the Poisson likelihood functions for treatment groups are:

$\text{event\_trt}_j \sim \text{Poisson}(\text{N\_trt}_j * p_j)$
$\text{event\_ctrl}_j \sim \text{Poisson}(\text{N\_ctrl}_j * q_j)$

With the same assumptions, the log odds ratio can be written as the form of the equation [5] too.

3.4 Prior Distributions of Risks for Binomial and Poisson likelihood

If the assumption of exchangeability hold, θ_j and ϕ_j: $j=1,\ldots,n$, are random-effects parameters that are drawn from some common prior distribution:

- $\pi(\theta_j) \sim \text{normal}(\mu_\theta, \sigma^2_\theta)$
- $\pi(\phi_j) \sim \text{normal}(\mu_\phi, \sigma^2_\phi)$

Since all the studies were randomized in this meta-analysis, the treatment and control can assume to have the same priority. In this meta-analysis, the following noninformative priors are placed on the hyperparameters μ, and σ^2:

- $\pi(\mu_\theta)$ and $\pi(\mu_\phi) \sim$ uniform(-1.5, 1.5)
- $\pi(\sigma_\theta^2)$ and $\pi(\sigma_\phi^2) \sim$ uniform(0, 8).

4. Meta Analysis Using Peto Method - A Frequentist's Approach

As a comparison, Peto's method (Yusuf S, Peto R, Lewis J, Collins R, Sleight P et al. 1985 {4}) was applied to estimate the pooled odds ratio. Peto and colleagues presented a method for pooling odds ratios. This method is not mathematically equal to the classical odds ratio, but it has come to be known as the 'Peto odds ratio'. The Peto odds ratio can cause bias, especially when there is a substantial difference between the treatment and control group sizes, but it performs well in many situations.

Assuming that we want to estimate $\theta_1 = \ldots = \theta_r = \theta$ Under a broad assumption one can derive: $\sum \hat{\theta}_i w_i \sim N(\theta \sum w_i, \sum w_i)$

Then it is possible to estimate θ by .

$$U = \hat{\theta} = \frac{\sum \hat{\theta}_i w_i}{\sum w_i} \qquad [6]$$

With 95% CI.

$$\hat{\theta} \pm 1.96 \sqrt{\frac{1}{\sum w_i}}$$

In order to test for the heterogeneity among the studies, the following Q statistics will be used:

$$Q = \sum_{i=1}^{r} (\hat{\theta}_i - \hat{\theta})^2 w_i = \sum \hat{\theta}_i^{\,2} w_i - U \sim \chi^2_{r-1} \qquad [7]$$

Q is a statistic to test heterogeneity among studies.

5. Results from Meta-Analyses and Model Validation

5.1 Results from Meta-Analyses

Table 1 shows the results of the meta analyses. The results from the Bayesian hierarchical modeling using Normal and Binomial Likelihood provided the following estimates. The pooled estimate of the odds ratio of oral candidiasis was 2.45 and 2.35 from the Normal and Binomial model respectively, indicating that oral candidiasis was more likely in inhaled corticosteroid (ICS) treatment compared to it in the placebo. The HPD credible interval estimates also confirmed this finding. The 95% credible interval are (1.50, 3.48) and (1.70, 3.08), from the Normal and Binomial model respectively. The Normal likelihood model has a higher variation among the treatment effects than that comes from the Binomial model. The sampling variance cannot be calculated for 36 studies by using the Normal likelihood.

As a comparison, Peto's method was applied to estimate the pooled odds ratio. The pooled estimate of the odds ratio of oral candidiasis was 2.72, indicating that oral candidiasis was more likely in an inhaled corticosteroid (ICS) treatment compared to that

in the placebo. The 95% confidence interval (CI) was (2.17, 3.42). Since this interval does not include 1, the oral candidiasis was statistically significantly worse in the ICS treatment compared to that in the placebo. The p-value from the z statistics was <0.0001, This confirms the same conclusion from the 95% CI. The Q statistic was 12.1 (p>0.9999), this means there was no evidence of heterogeneity among studies.

Table 1: Results from Meta-Analyses

Statistics	**Bayesian Hierarchical Modeling**		**Frequentist Approach**
	Normal approximation to binomial model	**Binomial Likelihood**	**Peto's Method**
Number of studies used in the analysis	14 of 50	50 of 50	50 of 50
Estimate of odds ratio	2.45	2.35	2.72
Standard deviation	0.52	0.37	0.32
95% HPD credible interval	(1.50, 3.48)	(1.70, 3.08)	
95% confidence interval			(2.17, 3.42)
Z statistics (p-value)			8.6 (<0.0001)
Q statistics (p-value)			12.1 (>0.9999)

5.2 Model Validation

One important assumption of Bayesian meta-analyses is exchangeability. Exchangeability means that no prior information suggests that any particular study is different from any other study. If the exchangeability assumption holds the Bayesian meta-analysis can be treated as a hierarchical model. A sequence of random quantities is said to be exchangeable if

$p(x_1,...,x_n) = p(x_{\pi(1)},...,x_{\pi(n)})$ holds for any subset permutation of $\pi(1,...,n)$

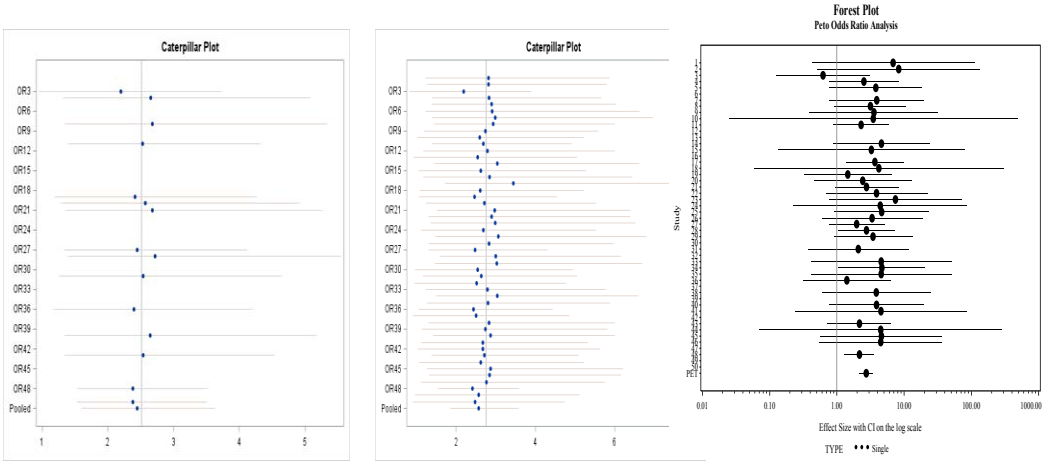
In clinical statistics, exchangeability {5} means that no prior information suggests that any particular study is different from any other study. There is no easy way to assess exchangeability; however, we can test its frequentist counterparts homogeneity and heterogeneity.

The Q heterogeneity statistics from Peto's method [7] is an important factor that needs to be assessed. If the Q heterogeneity statistics is not statistically significant, it indicates that there is no evidence of heterogeneity among studies.

Homogeneity can also be assessed by examining the forest plot created from a meta-analysis. The forest plots created by applying the Normal and Bionomial Bayesian hierarchical models and that created by applying the Peto's method are shown in figure 1.

Figure 1 Forest Plot

Bayesian Hierarchical Modeling		Frequentist's Approach
Normal approximation to binomial Model	Binomial Likelihood	Peto's Method

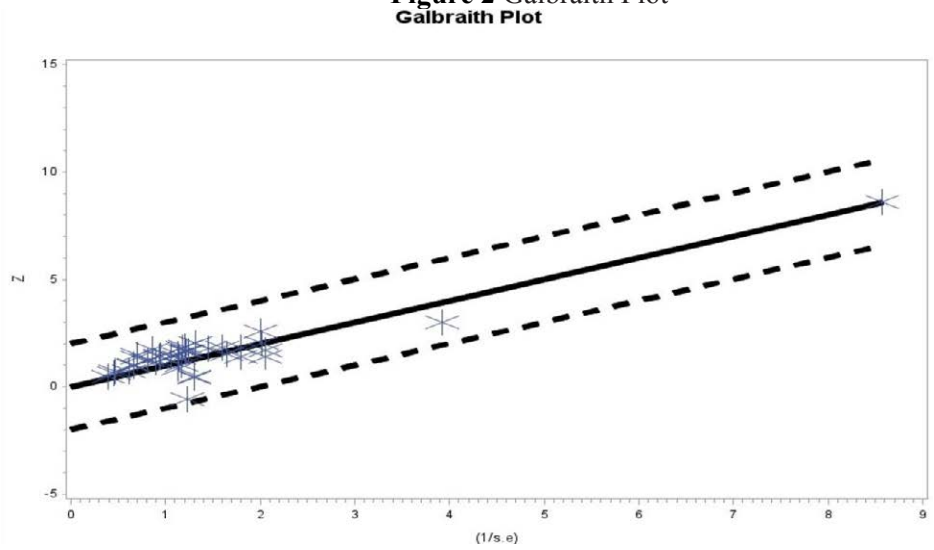


Homogeneity can also be examined with Galbraith plots. The Galbraith plot {6} is a scatter plot of standardized estimates (z-statistics divided by its standard error (SE)) against 1/SE. In this plot, the center line represents the pooled effect. i.e.

effect/se = (pooled effect) × 1/se

The 95% limits are 2 units above and below this line. We expect 95% of the points to be between these limits if there is no heterogeneity. The Galbraith plot that created by applying the Peto's Method are shown in figure 2.

Figure 2 Galbraith Plot



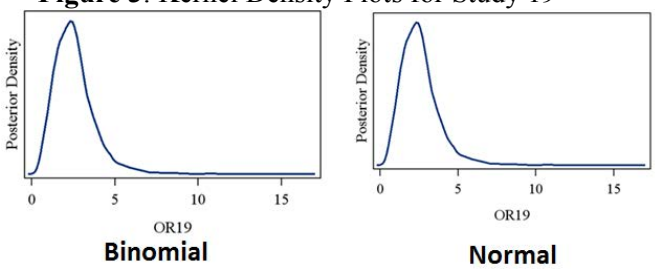
6. Discussion

The standard normal approximation method might not be appropriate because the approximation becomes less precise in extreme probabilities. For example, in the multi-studies dataset, the oral candidiasis rates were zero for many placebo and treatment groups. The odds ratio and its 95% credible intervals cannot be estimated in these cases so the pooled odds ratio was based only on the data from 14 of the 50 studies. Therefore, an alternative approach is to use the exact binomial likelihood approach as opposed to the normal approximation.

As a comparison, the kernel densities of the odds ratios for one of the 50 studies (study 19) are calculated by using the approximate normal likelihood and the exact binomial likelihood functions. Figure 3 compares the kernel density plots of the odds ratios that are produced by using these two likelihood functions for the Study 19. For this study, where the normal likelihood can be applied, the kernel density plots that are produced by using the approximate normal likelihood and the exact binomial likelihood are very similar.

In contrast, there are a huge number of studies for which the odds ratio cannot be estimated by the normal method. As described earlier, the pooled odds ratio was based only on the data from the 14 studies for which Y_j and s_j^2 can be calculated, whereas the exact binomial's estimate were based on 50 studies. Thus, there was a huge difference in the pooled odds ratio estimated from these two methods.

Figure 3: Kernel Density Plots for Study 19



The estimate on the pooled odds ratio from the exact binomial Bayesian method and the frequentist's Peto's method are very similar. The 95% credible interval and the 95% confidence interval are similar and consistent.

The exchangeability assumption for the Bayesian method was confirmed by Q heterogeneity statistic 12.1 ($p>0.9999$) from Peto's method and the forest plots and the Galbraith plots.

6. Conclusion

Results from this meta-analysis suggest that when the event rate is low, problematic data sets may be produced that result in incorrect or unstable estimates when analyzed. Similar findings were observed from Monica Bennett's research in 2013. {7} This was especially the case when the meta-analysis of odds ratios included studies with zero events observed.

Results from this meta-analysis also suggest that before performing meta-analysis by a Bayesian method, one should perform some data checks by the traditional frequentist method. The exchangeability assumption of the Bayesian method can be verified by assessing its frequentist counterparts homogeneity or heterogeneity. This will make sure the Bayesian analysis starts from a proper beginning. The possibility of extreme probabilities should be determined when deciding which likelihood distribution should be selected for the Bayesian meta-analysis. One method to do this assessment is to create caterpillar plots for the odds ratio and its 95% credible intervals.

References

{1} Hastings, W. K. (1970), "Monte Carlo Sampling Methods Using Markov Chains and Their Applications," Biometrika, 57, 97–109.

{2} Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953), "Equation of State Calculations by Fast Computing Machines," Journal of Chemical Physics, 21, 1087–1092.

{3} Spiegelhalter, D. J., Abrams, K. R., and Myles, J. P. (2004), "Bayesian Approaches to Clinical Trials and Health-Care Evaluation," Chichester, England: John Wiley & Sons.

{4} Yusuf S, Peto R, Lewis J, Collins R, Sleight P et al. "Beta blockade during and after myocardial infarction: an overview of the randomised trials," Progress in Cardiovascular Diseases 1985; 27:335–371.

{5} Jose B. "The Concept of Exchangeability and its Applications," Bhattacharya Memorial Volume, to be published in the Far East J. Math. Sci. December 9, 1996

{6} Galbraith, R.F. (1994), "Some Applications of Radial Plots," Journal of the American Statistical Association, 89 (428), pp. 1232–1242.

{7} Monica Bennett, Brenda C, Karen P, et al, "Comparison of Bayesian and Frequentist Meta-Analytical Approaches for Analyzing Time to Event Data", Journal of Biopharmaceutical Statistics, 23: 129–145, 2013