Bayesian Hierarchical Meta-Regression Using Study-Informed Priors

Jeff Gill
Division of Health Sciences, Department of Surgery
School of Medicine
Washington University, St. Louis

ABSTRACT

In this work provide a new approach to Bayesian meta-regression modeling in a hierarchical form with an enhanced ability to add prior information for model choice and coefficients. Our model includes more intra-study information than conventional meta-analysis and current forms of meta-regression. A key feature is simultaneous posterior estimates of standard meta-analysis effect summaries, such as log-relative risks, and model-averaged covariate coefficient estimates. We apply this new technology to....

1 Overview

Widespread policy and prescriptive procedures should be based on a body of studies rather than one single experiment or trial. This is not controversial. However, the manner in which many meta-analyses are performed invites severe criticism in the general statistics literature ( ) and elsewhere (Egger and Smith 1995)). Bluntly put, the heterogeneity of studies included in a meta-analysis is typically understated, and to the detriment of subsequent conclusions. Approaches that recognize unwanted, but present, heterogeneity are more robust describers of the studies effects. The most flexible and principled approach is the fully Bayesian approach of assuming that all unknown quantities should be described with probability statements, and that these probability statements can be parameterized.

In this work we revisit distributional assumptions underlying meta-analysis as normally practiced, and propose an new variant on the Bayesian approach to meta-regression that is more inherently robust than the current version, and demonstrate the inclusion of additional study information provides superior meta-results. The key distinction here is that by moving to a fully-specified meta-regression context, we include the individual covariate effects from the included model rather than just the treatment effect in isolation. Thus our contribution is the joint estimation covariate and treatment effects with priors that are informed by individual model quality as well as qualitatively informed views that can come from previous research, clinical information, or additional model features from these studies.

2 Basics

There are essentially three distributional approaches to meta-analysis. The so-called *fixed effects* approach assumes that the underlying quantity of interest (parameter estimates, relative risk, odds ratio, or general effect size) is fixed by nature and each study included in the analysis simply draws from a sampling distribution around this fixed point. Therefore the interpretation of the summarized results includes the assumption that the square-root of the variance of found quantities in the studies is the standard error of a mean estimate of the true fixed value. The *random effects* approach assumes (Higgins *et al.* 2009) that there is sufficient heterogeneity and uncertainty for the quantity of interest that it really defines a distribution rather than a fixed point in the complete population sense. Therefore study observations of this quantity are simply draws from this unknown distribution. The Bayesian approach uses random effects perspective as a starting point and considers parameterizations of the distribution of random effects rather than just assuming a basic (usual normal) form. Therefore we have escalating levels of importance of the distributional assumption.

Despite routine tests that are made, the fixed effects approach is rarely justified. This is because included studies almost always have different mixes of right-hand side covariates in the original models. It is well-appreciated that minor differences in regression-style GLM specifications can yield immense differences in inference and prediction. Therefore it is a very strict assumption

to say that the studies produce IID draws from a fixed sampling distribution.

The Random effect approach is much more principled since it admits a level of heterogeneity between the included studies such that the quantity of interest needs to be modeled distributionally. That is, there may be a modal region for strong effects, but differing specifications and sampling strategies cause deviations from this point of centrality. The Bayesian approach improves-upon the random effects approach because it attempts to *model* the underlying cause for study heterogeneity.

The key distinction between meta-analysis and meta-regression is the inclusion of the model components in addition to the treatment/control contrast. Therefore each included study contributes covariate information in addition to the key variable of interest. The challenge here, of course, is that each study will specify a different mix of right-hand-side explanatory variables and the meta-regression needs to account for this additional heterogeneity. Our approach is to consider each of these models as a selection from the super-set of possible regression parameters defined by the union of explanatory variables across the included models.

3 Bayesian Meta-Analysis

The Bayesian approach to meta-analysis is increasingly popular since it explicitly considers distributional properties of unknown treatment effects (Higgins and Spiegelhalter 2002). Consider a set of K studies, each of which provide an estimated treatment effect, $\hat{\theta}_1, \cdots, \hat{\theta}_K$. This treatment effect is often: a regression parameter estimate, a log-odds ratio, log relative risk, or absolute relative risk. Under very general circumstances each of these $\hat{\theta_k}$ can be considered independently normally distributed by the asymptotic theories:

$$\hat{\theta}_k \sim f(\hat{\theta}_k|\theta_k, \sigma_k^2) = N(\theta_k, \sigma_k^2), \tag{1}$$

where $\theta_1, \cdots, \theta_K$ are true individual study treatment effects and $\sigma_1^2, \cdots, \sigma_K^2$ are the individual study sample variance. Under the random effects model (DerSimonian and Laird 1986), it is assumed that σ_k^2 equals $\hat{\sigma}_k^2$, standard error estimate of θ_k in study k and

$$\theta_k \sim f(\theta_k|\theta, \sigma^2) = N(\theta, \sigma^2), \tag{2}$$

where θ and σ^2 are the true mean and the variance of population distribution. In a Bayesian meta-analysis, we assume that θ, σ^2 have a prior distribution $\pi(\theta, \sigma^2)$. Then the posterior distribution of θ can be written as

$$p(\theta|\hat{\theta}_1, \cdots, \hat{\theta}_K) \propto \int f(\hat{\theta}_k|\theta_k, \sigma_k^2)\{\prod_k \pi(\theta_k|\theta, \sigma^2)\}\pi(\theta, \sigma^2)d\theta_1 \cdots d\theta_K d\sigma^2. \tag{3}$$

This traditional Bayesian meta-analysis method considers only the treatment effect, θ.

In many cases, the treatment effects are estimated by fitting the generalized linear model (GLM) including linear regression and logistic models. To adjust confounding factors such as age and

gender, authors tend to include the confounding factors in the GLM equation. We now propose a Bayesian meta-analysis approach exploiting the covariate parameter estimates. We call the confounding factor the covariate. To account for the different selection of covariates in each study (model), we define X_k as a set of covariates in a model k assumed in a study k. For example, X_k can be (gender, age). We define a union of X_k across studies as X. We write $X = (x_1, \cdots, x_J)$ where $x_1, \cdots, x_J$ are distinct covariates in X. We define an index set $I_k = \{j; x_j \in X_k\}$ and write $X_k = \{x_{jk}, j \in I_k\}$. Consider a full model in which in the RHS of the model equation is a function of $X\gamma$ where $\gamma = (\gamma_1, \cdots, \gamma_J)^T$ is a parameter vector. Let γ_{jk} denote a parameter corresponding to x_{jk}. If the linear regression model is assumed, then the model equation of the study k can be written as

$$y = \alpha + \theta_k t + \sum_{j \in I_k} \gamma_{jk} x_{jk} + \epsilon. \tag{4}$$

Let $\hat{\gamma}_{jk}$ be the estimate for γ_{jk} obtained by fitting the model above. Define $\phi_k = (\theta_k, \gamma_{jk}, j \in I_k)$ and $\hat{\phi}_k = (\hat{\theta}_k, \hat{\gamma}_{jk}, j \in I_k)$. Under the normality assumption of ϵ, $\hat{\phi}_k$ is normally distributed as

$$\hat{\phi}_k \sim f(\hat{\phi}_k | \phi_k, \Sigma_k) = N(\phi_k, \Sigma_k). \tag{5}$$

We assume that Σ_k equals $\hat{\Sigma}_k$ and that ϕ_k is a random draw from a pdf

$$\int \pi(\phi_k, \{\gamma_{jk}, j \notin I_k\} | \phi = (\theta, \gamma), \Sigma) d\gamma_{(jk, j \notin I_k)}. \tag{6}$$

Suppose that (ϕ, Σ) has a prior distribution $\pi(\phi, \Sigma)$. Then, the posterior distribution given all estimates from K studies is

$$p(\theta, \gamma | \hat{\phi}_1, \cdots, \hat{\phi}_K) \propto \int f(\hat{\phi}_k | \phi_k, \Sigma_k) \{\prod_k \int \pi(\phi_k, \{\gamma_{jk}, j \notin I_k\} | \phi, \Sigma) d\gamma_{(jk, j \notin I_k)}\} \pi(\phi, \Sigma) d\phi_1 \cdots d\phi_K d\Sigma. \tag{7}$$

4 Measuring Heterogeneity

Heterogeneity is a function of the differences between study design, data, and modeling approach for a set of studies in some meta-analysis. Bayesian and non-Bayesian approaches both need to consider whether the estimated level of heterogeneity makes the studies suspect as a cohesive group. Since collection of the included studies is the most time-consuming and controversial component of any meta-analysis project, this is an important consideration. Presumed or suspected heterogeneity between studies is the primary reason to move from a fixed effects to a random effects model. Here we are concerned with whether the inevitable differences between included studies exceeds a subjective level. Usually this level is a function of variance, although there are different ways of measuring it (Petitti 2001). Often the level of heterogeneity is measured with the Q-statistic

(DerSimonian and Laird 1986), which is the sum of squares of the treatment effect about the common mean weighted by the inverse of the estimated mean in the study: $Q = \omega_k(\delta_k - \bar{\delta})^2$, where $\bar{\delta} = \frac{1}{K}\sum \omega_k \delta_k$. Under the null hypothesis of homogeneity between studies this is distributed χ^2 with $K-1$ degrees of freedom. Note that this assumes a common sampling variance such that $\omega_k = 1/s_k^2$ is considered to be a draw from the same distribution, although there is some controversy about the efficacy of this assumption. The Q-statistic sampling variance from the kth study is defined for three effects of interest in Table 1.

Table 1: Q-Statistic For Different Treatment Models

Log-Odds Ratio:	$s_k^2 = \frac{1}{r_k^C n_k^C} - \frac{1}{n_K^C} + \frac{1}{r_k^T n_k^T} - \frac{1}{n_K^T}$
Absolute Relative Risk:	$s_k^2 = \frac{r_k^C(1-r_k^C)}{n_k^C} + \frac{r_T^C(1-r_k^T)}{n_k^T}$
Log Relative Risk:	$s_k^2 = \frac{1}{n_k^C r_k^C(1-r_k^C)} + \frac{1}{n_k^T r_k^T(1-r_k^T)}$

It is well-known that the Q-statistic provides a low power test for small K, and is also overly sensitive to minor heterogeneity for large K. On alternative assuming a random effects model is τ^2, the between-study variance measured in units of the outcome variable. Since this is measured particularistically on the study outcome, often τ^2 values cannot be compared between meta-analyses. To make up for this deficiency, Higgins and Thompson (2002) created the I^2 statistic, which is the percent of total between variance of the effect sizes due to true heterogeneity $I^2 = 100\frac{Q-(K-1)}{Q}$ (for $Q > 0$, otherwise zero). This is interpreted as the percent of total variability that is due to between-study variance. Obviously τ^2 and I^2 are proportionally related.

5 Describing the Sampler For Estimation

6 Application To (Something Cancer Oriented)

7 Conclusions

8 References

Bohlius F, Schmidlin K, Brillant C, Schwarzer G, Trelle S, Seidenfeld J, Zwahlen M, Clarke M, Weingart O, Kluge S, Piper M, Rades D, Steensma D P, Djulbegovic B, Fey M F, Ray-Coquard I, Machtay M, Moebus V, Thomas G, Untch M, Schumacher M, Egger M, Engert A. Recombinant human erythropoiesis-stimulating agents and mortality in patients with cancer: a meta-analysis of randomised trials. *Lancet* 2009; **373**; 1532-1542.

Chib S. Marginal likelihood from the Gibbs output. *Journal of the American Statistical Association* 1995; **90**; 1313-1321.

Chib S, Jeliazkov, I. Marginal Likelihood from the Metropolis-Hastings Output. *Journal of the American Statistical Association* 2001; **96**; 270-281.

DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clinical Trials* 1986; **7**; 17788.

Egger M, Smith G. Misleading meta-analysis. *British Medical Journal* 1995; **6982**; 752-754.

Gill, J. *Bayesian Methods For the Social and Behavioral Sciences.* New York: Chapman & Hall/CRC; 2007.

Harville D A. Maximum likelihood approaches to variance components estimation and to related problems. *Journal of the American Statistical Association* 1977; **72**: 320-340.

Hayden J, van Tulder M W, Tomlinson G. Systematic review: strategies for using exercise therapy to improve outcomes in chronic low back pain. *Annals of internal medicine* 2005; **142**; 776-785.

Higgins J P T, Spiegelhalter D J. Being sceptical about meta-analyses: a Bayesian perspective on magnesium trials in myocardial infarction. *International Journal of Epidemiology* 2002; **31**; 96-104.

Higgins J P T, Thompson S G. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* 2002; **11**; 1539-1558.

Higgins J P T, Thompson S G, Spiegelhalter D J. A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society. Series A* 2009; **172**; 137-159.

Laird N M, Ware J H. Random-effects models for longitudinal data. *Biometrics* 1982; **38**; 963974.

Lopes H F, Müller P, Rosner G L. Bayesian meta-analysis for longitudinal data models using multivariate mixture priors. *Biometrics* 2003; **59**; 66-75.

Petitti, D. Approaches to heterogeneity in meta-analysis. *Statistics in Medicine* 2001; **20**; 3625-3633.

Sterne J A C, Egger M. Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis. *Journal of Clinical Epidemiology* 2001; **54**; 10461055.

Thompson S G, Higgins J P T. How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine* 2002; **21**; 1559-1573.

Warn, Thompson, and Spiegelhalter. Bayesian random effects meta-analysis of trials with binary outcomes: methods for the absolute risk difference and relative risk scales *Statistics in Medicine* 2002; **21**; 1601-1623.