STATISTICS IN MEDICINE, VOL. 15, 2733–2749 (1996)

BORROWING STRENGTH FROM EXTERNAL TRIALS IN A META-ANALYSIS

JULIAN P. T. HIGGINS*

Department of Applied Statistics, The University of Reading, P.O. Box 240, Reading, RG6 6FN, U.K.

AND

ANNE WHITEHEAD

Medical and Pharmaceutical Statistics Research Unit, The University of Reading, P.O. Box 240, Reading, RG6 6FN, U.K.

SUMMARY

There exists a variety of situations in which a random effects meta-analysis might be undertaken using a small number of clinical trials. A problem associated with small meta-analyses is estimating the heterogeneity between trials. To overcome this problem, information from other related studies may be incorporated into the meta-analysis. A Bayesian approach to this problem is presented using data from previous meta-analyses in the same therapeutic area to formulate a prior distribution for the heterogeneity. The treatment difference parameters are given non-informative priors. Further, related trials which compare one or other of the treatments of interest with a common third treatment are included in the model to improve inference on both the heterogeneity and the treatment difference. Two approaches to estimating relative efficacy are considered, namely a general parametric approach and a method explicit to binary data. The methodology is illustrated using data from 26 clinical trials which investigate the prevention of cirrhosis using beta-blockers and sclerotherapy. Both sources of external information lead to more precise posterior distributions for all parameters, in particular that representing heterogeneity.

1. INTRODUCTION

A random effects meta-analysis is an increasingly popular method of summarizing the results of a set of related randomized clinical trials when there is unexplained heterogeneity between them. Methodology for combining simple estimates of the relative efficacy of two treatments is given, for example, by DerSimonian and Laird[1] or Whitehead and Whitehead.[2] Data of a specific type may be modelled explicitly.[3] Similar approaches can be taken using Bayesian methodology.[4,5] However, irrespective of the approach taken, a reasonable number of clinical trials is required to give results with sufficient precision.

There exists a variety of situations in which a meta-analysis of a small number of clinical trials might be undertaken. This would be the case in a regulatory submission of a new drug, in the early stages of a cumulative meta-analysis, or if the condition being treated is rare. In many cases there is information from sources other than the trials under consideration which is relevant to

*Current address: Department of Epidemiology and Biostatistics, Institute of Child Health, 30 Guildford Street, London WC1N 1EH, U.K.

CCC 0277–6715/96/242733–17

the treatment comparison of interest. Such information may be incorporated into the meta-analysis to improve precision in the estimation of both heterogeneity and relative efficacy.

The traditional Bayesian approach is to formulate prior distributions which represent subjective opinion. By contrast, this paper describes the incorporation of real data from previous studies using Bayesian techniques. The first source, historical meta-analyses of clinical trials in the same therapeutic area, improves estimation of the degree of heterogeneity between the trials. The second source is clinical trials investigating the efficacy of one or other of the treatments of concern relative to some common third treatment. This follows the ideas of Begg and Pilote[6] who make use of uncontrolled studies to assess treatment effects in inference concerning treatment differences in a small set of controlled studies.

This paper discusses methods for performing a random effects meta-analysis of binary data. In Section 2 two approaches to the estimation of relative efficacy are reviewed. The first, based on Whitehead and Whitehead,[2] is a general parametric approach, the second, as described by Smith,[7] is specific to binary data. The chosen method in practice would depend on information which was available from the individual studies. When patient numbers and success rates are available the exact binomial approach is the more suitable. However, the general parametric approach requires only two summary statistics from each trial, an estimate of relative efficacy and an estimate of its variance. The log odds ratio is considered here as the measure of relative efficacy for binary data. Both the maximum likelihood estimate and an approximate maximum likelihood estimate based on efficient score and Fisher's information statistics will be considered.

The following section presents a Bayesian approach, with analyses being performed using the Gibbs sampler. Incorporation of information from historical meta-analyses is described in Section 4.

The inclusion of related trials involving a third treatment is described in Section 5. Some trials in the main meta-analysis data set may compare all three treatments. Assumptions must be made about the heterogeneity parameters across the various treatment comparisons, leading to a covariance term in the model. In addition, two study-level estimates of different treatment comparisons are not, in general, independent. An estimate of their covariance is derived.

The methodology is illustrated in Section 6 using data from a set of clinical trials which investigate the prevention of first bleeding in cirrhosis using beta-blockers and sclerotherapy.[8] This leads to a discussion of the relative merits of the two approaches to estimation.

2. RANDOM EFFECTS META-ANALYSIS

Suppose we have k related trials on which a meta-analysis is to be performed to compare two treatments A and B, and suppose that underlying the ith trial is an unknown relative efficacy θ_i. In the case of binary data relative efficacy might be expressed as a log odds ratio of success on A relative to B. If these θ_i can be assumed to be the same across all trials, then a fixed effect analysis is appropriate. If there is evidence that the θ_i are not the same, then an explanation of this should be sought. Heterogeneity may exist for a variety of reasons: there may be differences in execution of the trials or in patient populations, or the meta-analysis may be of trials investigating different (but still related) treatments. If heterogeneity is found which cannot be explained, then it may be accounted for in a random effects analysis. It is generally assumed that the underlying relative efficacy for the population in the ith trial is drawn from a normal distribution with some mean, μ, representing the overall average relative efficacy and a variance parameter, τ^2, representing the heterogeneity between the trials.[1]

2.1. A model for a meta-analysis of binary data

Suppose each patient in the ith trial has probability π_{Ai} of a successful outcome on treatment A or probability π_{Bi} of a successful outcome on treatment B. Let s_{Ai} and s_{Bi} be the numbers of successful patients on treatments A and B, respectively, in that trial, out of total numbers of patients n_{Ai} and n_{Bi}. The basic model for the meta-analysis is then, for $i = 1, \ldots, k$,

$$s_{Ai} \sim \text{Bin}(n_{Ai}, \pi_{Ai}) \tag{1}$$

$$s_{Bi} \sim \text{Bin}(n_{Bi}, \pi_{Bi}) \tag{2}$$

$$\log \frac{\pi_{Ai}}{(1 - \pi_{Ai})} = \phi_i \tag{3}$$

$$\phi_i - \log \frac{\pi_{Bi}}{(1 - \pi_{Bi})} = \theta_i \tag{4}$$

$$\theta_i \sim \text{N}(\mu, \tau^2). \tag{5}$$

The ϕ_i are fixed nuisance parameters in a frequentist approach.

2.2. Methods of estimation

This section describes two approaches to estimating parameters of interest in a meta-analysis, and frequentist methods for inference about the overall relative efficacy μ and the heterogeneity parameter τ^2. The first, a general parametric approach,[2] estimates the log odds ratios, θ_i, separately from each trial and combines these to form estimates of μ and τ^2. The second makes inference about the probabilities π_{Ai} and π_{Bi} in conjuction with the θ_i, μ and τ^2 in an explicitly binomial model.

2.2.1. A general parametric approach

A general parametric method for a random effects analysis requires an estimate of relative efficacy and an estimate of its variance from each trial. Let these be denoted by y_i and v_i, respectively, for the ith trial. We desire estimates for which the assumption $y_i \sim \text{N}(\theta_i, v_i)$ can be made approximately. It is henceforward assumed that v_i is known.

The likelihood function for the unknown parameters is

$$L(\mu, \tau^2, \theta_i \mid y_i, v_i) = \prod_{i=1}^{k} \frac{1}{\sqrt{(2\pi v_i)}} \exp\left(-\frac{1}{2v_i} (y_i - \theta_i)^2 \right) \frac{1}{\sqrt{(2\pi\tau^2)}} \exp\left(-\frac{1}{2\tau^2} (\theta_i - \mu)^2 \right). \tag{6}$$

There are two common estimators of a log odds ratio in a clinical trial which yield values for y_i and v_i. The maximum likelihood estimate is

$$y_{i,\text{MLE}} = \log \left(\frac{s_{Ai}(n_{Bi} - s_{Bi})}{s_{Bi}(n_{Ai} - s_{Ai})} \right) \tag{7}$$

with approximate variance

$$v_{i,\text{MLE}} = \frac{1}{s_{Ai}} + \frac{1}{s_{Bi}} + \frac{1}{n_{Ai} - s_{Ai}} + \frac{1}{n_{Bi} - s_{Bi}}.$$

In the case of a treatment group having no successes or no failures, a correction factor of a half may be added to each of $s_{Ai}, s_{Bi}, n_{Ai} - s_{Ai}$ and $n_{Bi} - s_{Bi}$.

An alternative estimate is an approximate maximum likelihood estimate based on efficient score and Fisher's information statistics. It has the advantage that these statistics can readily be calculated for most types of data. Statistics for survival, ordered categorical and normally distributed data are given in Whitehead.[9] Let Z_i and V_i represent the efficient score and Fisher's information for the measure of treatment difference, in our case the log odds ratio, respectively. Then asymptotically and for small treatment differences,

$$Z_i \sim \mathrm{N}(\theta_i V_i, V_i). \tag{8}$$

Thus an estimate of relative efficacy is given by

$$y_i = Z_i/V_i \tag{9}$$

with approximate variance $v_i = 1/V_i$.

The statistics for the log odds ratio of success on treatment A relative to treatment B are

$$Z_i = \frac{s_{\mathrm{A}i} n_{\mathrm{B}i} - n_{\mathrm{A}i} s_{\mathrm{B}i}}{n_{\mathrm{A}i} + n_{\mathrm{B}i}} \tag{10}$$

$$V_i = \frac{n_{\mathrm{A}i} n_{\mathrm{B}i} (s_{\mathrm{A}i} + s_{\mathrm{B}i})(n_{\mathrm{A}i} + n_{\mathrm{B}i} - s_{\mathrm{A}i} - s_{\mathrm{B}i})}{(n_{\mathrm{A}i} + n_{\mathrm{B}i})^3}. \tag{11}$$

Note that Z_i can be expressed as $s_{\mathrm{A}i} - n_{\mathrm{A}i}(s_{\mathrm{A}i} + s_{\mathrm{B}i})/(n_{\mathrm{A}i} + n_{\mathrm{B}i})$; so is the Peto $O - E$ statistic.[10] The denominator of $(n_{\mathrm{A}i} + n_{\mathrm{B}i})^3$ in V_i differs from that of $(n_{\mathrm{A}i} + n_{\mathrm{B}i})^2(n_{\mathrm{A}i} + n_{\mathrm{B}i} - 1)$ in other papers[2,10] since it is based on the assumption of binomially distributed rather than hypergeometrically distributed responses. In practice the difference between the two is small.

A disadvantage of the Z/V method is that large treatment differences can be severely underestimated. In fact, with binary data, simulations undertaken by the authors have shown this to be the case for true log odds ratios greater than one, irrespective of the numbers of patients in the trial.

Given a point estimate, $\hat{\tau}^2$, of the heterogeneity parameter, a frequentist meta-analysis may be performed based on the general parametric approach. A common, and simple, estimate of τ^2 is a method of moments estimate (see, for example, DerSimonian and Laird[1]). An estimate of μ is given by the weighted average of the y_i using their inverse variances $(v_i + \hat{\tau}^2)^{-1}$ as weights. The estimate has approximate variance $(\sum_{i=1}^{k} (v_i + \hat{\tau}^2)^{-1})^{-1}$. A more complicated, but more accurate, likelihood approach is described by Hardy and Thompson.[11]

2.2.2. *An exact binomial approach*

An approach which bypasses the assumption of known variance v_i involves modelling the raw data (n's and s's) rather than estimates of log odds ratios. The likelihood of the parameters is

$$L(\mu, \tau^2, \theta_i \mid s_{\mathrm{A}i}, n_{\mathrm{A}i}, s_{\mathrm{B}i}, n_{\mathrm{B}i}) = \prod_{i=1}^{k} \binom{n_{\mathrm{A}i}}{s_{\mathrm{A}i}} \pi_{\mathrm{A}i}^{s_{\mathrm{A}i}} (1 - \pi_{\mathrm{A}i})^{n_{\mathrm{A}i} - s_{\mathrm{A}i}} \times \binom{n_{\mathrm{B}i}}{s_{\mathrm{B}i}} \pi_{\mathrm{B}i}^{s_{\mathrm{B}i}} (1 - \pi_{\mathrm{B}i})^{n_{\mathrm{B}i} - s_{\mathrm{B}i}} \frac{1}{\sqrt{(2\pi\tau^2)}} \exp\left(-\frac{1}{2\tau^2} (\theta_i - \mu)^2 \right). \tag{12}$$

Making inferences from this likelihood involves the nuisance parameters, ϕ_i, and so does not lead to as simple an analysis as the general parametric approach above. However, a frequentist random effects meta-analysis may be performed by logistic regression using, for example, the statistical package EGRET.[12] See Smith *et al.*[13] for details.

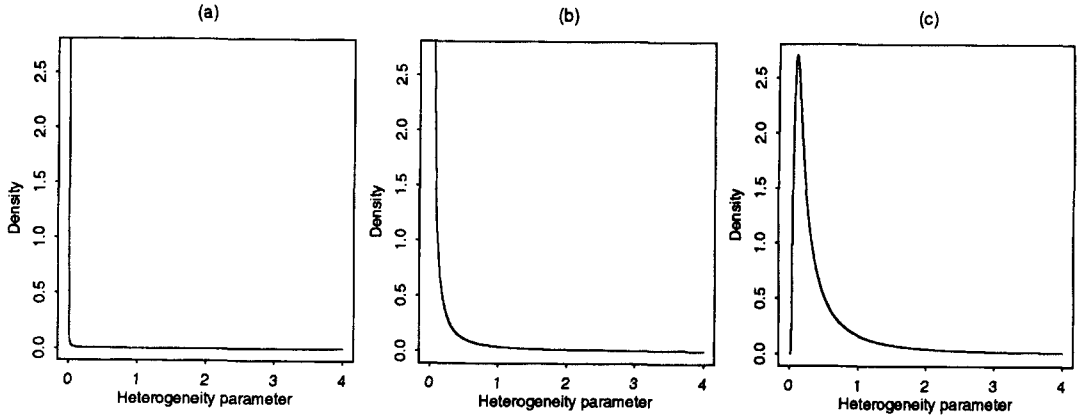

Figure 1. Prior distributions for τ^2. (a) IG($10^{-3}, 10^{-3}$); (b) IG(0.5, 0.005); (c) IG(1.0, 0.2).

3. A BAYESIAN META-ANALYSIS

The frequentist approaches above provide a simple analysis of a pairwise treatment comparison. However, they are restricted to such an analysis. The model may also be analysed within a Bayesian framework, and within that framework additional information may be incorporated into the analysis through prior distributions and by adding extra treatment comparisons to the model. A basic Bayesian approach extends (1)–(5) by incorporating prior distributions for the hyperparameters μ and τ^2. In Bayesian terminology (5) represents an assumption of exchangeability between the relative efficacies underlying the trials.

Conjugate priors for μ and τ^2 are normal and inverse gamma (IG), respectively. An IG(α, λ) distribution is specified here as $p(\tau^2) = \lambda^{\alpha}(\Gamma(\alpha))^{-1}(\tau^2)^{-\alpha-1}e^{-\lambda\tau^{-2}}$. non-informative priors may be approximated by a very flat normal distribution for μ and an inverse gamma distribution with near zero parameters for τ^2, for example N(0, 10^6) and IG($10^{-3}, 10^{-3}$).

The Jeffreys non-informative prior for the variance of a normal distribution[14] would be to take $p(\tau^2) \propto 1/\tau^2$, which corresponds to an IG(0, 0). For a discussion of why this is not appropriate in a random effects analysis see DuMouchel and Waternaux,[15] or the BUGS manual.[5] The vague prior IG($10^{-3}, 10^{-3}$) is illustrated in Figure 1(*a*).

Joint posterior distributions are produced by multiplying (6) by $p(\mu)p(\tau^2)$ or (12) by $p(\mu)p(\tau^2)\prod_i p(\phi_i)$, where $p(\phi_i)$ is taken here to be a non-informative N(0, 10^3). The former retains the assumption of known variances for the log odds ratio estimates, but the latter accounts for uncertainty surrounding all unknown parameters. In either case, integration to calculate marginal posterior distributions for the parameters cannot be performed analytically. However, the full conditional distribution of each parameter may be determined easily. The Gibbs sampler[16] can therefore be used to produce approximate samples from the posterior distributions of μ, τ^2 and the θ_i, and hence to form point estimates and posterior probability intervals. The program BUGS[5,17] is used here, with convergence checked using a mixture of the methods of Raftery and Lewis,[18] Geweke,[19] Heidelberger and Welch[20] and by observing the traces of posterior samples, implemented using CODA[21] (Convergence Diagnosis and Output Analysis Software for Gibbs sampling output).

4. AN EMPIRICAL PRIOR FOR HETEROGENEITY

In a random effects meta-analysis, heterogeneity may be considered as unexplainable variation between relative efficacies in different trials. It could be due, among other sources, to demographic factors or to different hospital procedures. One might also include small differences in protocols, too difficult to quantify and thus to introduce as covariates, as random variation.

Such random variation is commonly found, and a meta-analysis may be able to draw on results of historical meta-analyses. In particular, trials of similar interventions for similar treatments are likely to encounter the same reasons for random variation between relative efficacies, and may therefore be expected to produce similar estimates of the amount of heterogeneity. In the following, a prior distribution is formulated for τ^2 for use in the example of Section 6 based on 18 meta-analyses in the area of gastroenterology. All meta-analyses involved binary data and had similar clinical endpoints. A variety of methods of varying complexity may be used to formulate prior distributions. These are described in the remainder of this section.

Smith[7] formed a prior from 30 meta-analyses in a variety of medical areas. By considering the empirical cumulative distribution function (ECDF) and a kernel density estimate of 30 method of moments estimates of heterogeneity, she decided upon an IG(0·5, 0·005) distribution (Figure 1(*b*)). She found little difference between the use of this prior and the non-informative prior in a meta-analysis of 22 randomized trials which investigated the clinical benefits of selective decontamination of the digestive tract. The information contained in the 22 trials overwhelmed that contained in the prior distribution. In much smaller meta-analyses this will not be the case.

We present a less *ad hoc* method for calculating a data based prior. Our prior distribution results from the combination of all the data from the historical meta-analyses into one large Bayesian meta-analysis of meta-analyses. This treats the heterogeneity in the jth meta-analysis, τ_j^2, as a random effect from an IG(α, λ) distribution. Prior distributions are placed on the hyperparameters α and λ. The predictive distribution of a 'new' heterogeneity parameter, say τ^2_{new}, may be used to form a prior distribution for τ^2 in the present meta-analysis. This incorporates the uncertainty surrounding all parameters into the prior, but involves the approximation of the simulated posterior by a parametric distribution. Alternatively, summary statistics from the posterior distributions of α and λ may be used as parameters for a prior distribution.

An alternative, more restrictive, model is to assume the same heterogeneity in each of the historical meta-analyses. A non-informative prior for this τ^2 should be updated to a posterior distribution suitable for use as a prior distribution in the present meta-analysis.

4.1. An empirical prior distribution for heterogeneity in gastroenterology meta-analyses

After a Medline search and examination of over a hundred papers, ten published meta-analysis papers[22-31] were identified of studies in gastroenterology. They refer to 18 sets of very similar types of study with roughly the same primary endpoint of whether gastrointestinal bleeding (re)occurs after treatment. From these 18 meta-analysis data sets the method of moments estimate of τ^2 was calculated. The closest fitting inverse gamma distribution to the ECDF of these estimates of τ^2 was found to be an IG(1·0, 0·2) (Figure 1(*c*)). The same conclusion is reached if the ECDF is formed from posterior medians of τ^2 resulting from Bayesian analyses of the meta-analysis data sets using the model in Section 3. Having larger parameters than the distribution found by Smith,[7] this prior will be more influential.

An alternative prior distribution for τ^2 is formed by performing a meta-analysis of the meta-analyses. This assumes that the heterogeneity parameters in the 18 meta-analyses are a sample from an inverse gamma distribution with parameters α and λ. Prior distributions must be placed on the parameters α and λ. A gamma distribution provides a conjugate prior for λ,

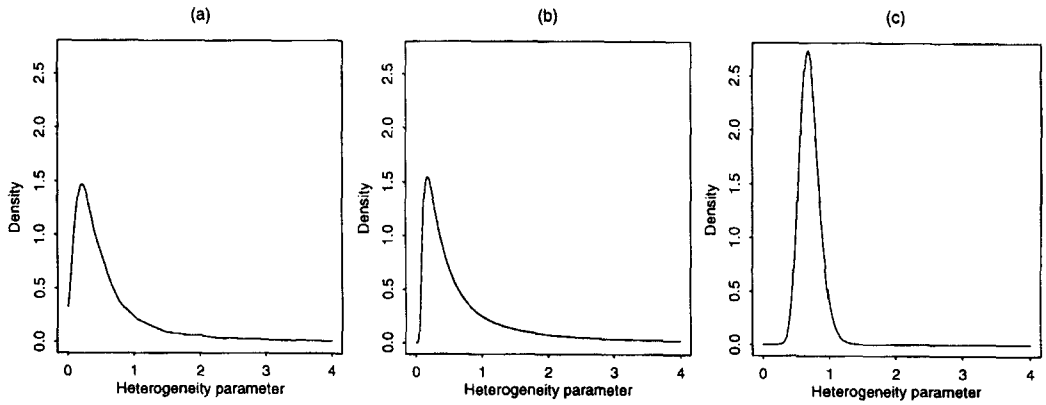

Figure 2. Kernel density estimates of posterior distributions following meta-analysis of 18 gastroenterology meta-analysis data sets: (*a*) assuming random effects for heterogeneity parameters, 15,000 iterations of the Gibbs sampler following a burn-in of 1000; (*b*) a parametric approximation to (*a*): IG(1·0, 0·35); (*c*) assuming equal heterogeneity parameter in all meta-analyses, 15,000 iterations following a burn-in of 1000

but not for α. However, a discrete approximation to a gamma distribution is possible in BUGS. This makes use of a categorical variable taking many categories with probabilities proportional to the required inverse gamma distribution (Nicky Best, personal communication). Vague gamma(0·001, 0·001) distributions with unit mean and a variance of 1000 were used as prior distributions for both parameters.

The prior distribution for τ^2 in the Pagliaro meta-analysis discussed in Section 6 may be based either on the posterior medians of α and λ or on the simulated predictive distribution of τ^2, that is, that of τ^2_{new}. A run of 15,000 iterations of the Gibbs sampler, following an initial period of 1000 iterations, gave posterior medians of 1·54 for α (95 per cent probability interval 0·49 to 3·8) and 0·52 for λ (95 per cent probability interval 0·05 to 2·0). The predictive distribution for τ^2 has posterior median 0·42 and a 95 per cent posterior probability interval from 0·05 to 7·1. The kernel density of the posterior distribution of τ^2_{new} is illustrated in Figure 2(*a*). A close-fitting inverse gamma distribution is found to be an IG(1·0, 0·35) (Figure 2(*b*)). This agrees fairly well with the IG(1·0, 0·2) distribution derived by the simpler method.

If the random effect is removed, implying the assumption that all τ_j^2's are equal, then, after 15,000 iterations following a burn-in of 1000, the posterior distribution of the common heterogeneity parameter has median 0·70 and 95 per cent probability interval 0·47 to 1·0. A kernel density estimate is shown in Figure 2(*c*). However, 14 of the individual method of moments estimates lie outside this interval. This reflects the uncertainty surrounding each estimate, being based on a small number of trials. The distribution also gives near-zero heterogeneity (that is, that a fixed effect meta-analysis model is a distinct possibility) extremely low probability. The authors believe a more conservative prior derived from the random effects approach to be more suitable.

5. INCLUDING TRIALS OF A THIRD TREATMENT

Many meta-analysis data sets contain information on more than two treatments. Often interest is in a new active treatment which has been compared sometimes with a placebo control and sometimes with a standard active treatment. Alternatively it may be that the two treatments in the

primary comparison have at some stage been compared with some common treatment. A traditional meta-analysis would make use of a subset of such data, for example new versus placebo, or new versus standard. A Bayesian meta-analysis allows two or more relative efficacies to be analysed simultaneously. Alternatively, a single treatment difference may be estimated more precisely by incorporating data from other treatment comparisons.

Given three treatments A, B and C, there exist three population pairwise comparisons, μ_{AB}, μ_{AC} and μ_{BC}, say. Each may be written in terms of the other two (see DuMouchel[32]), hence in practice only two of the comparisons need be included in a model.

It is assumed here that the heterogeneity parameter associated with the various relative efficacies is the same. This would require investigation in any particular data set, but the authors do not believe there will often be sufficient reason to reject the assumption. However, if one treatment is expected to have substantially more variable results between trials (perhaps a treatment with substantial geographical variability or heavily dependent on the patients involved) then the assumption may not be valid. The analysis may be carried out assuming different heterogeneity parameters for different treatment comparisons. This is slightly more complicated if one or more of the trials compares all three treatments. This scenario is discussed in Section 5.1.

In the simple, but unusual, case where no trial in the meta-analysis includes results on all three treatments, the model may be written as follows, where $i = 1, \ldots, k$:

$$s_{Ai} \sim Bin(n_{Ai}, \pi_{Ai}), \qquad \phi_{Ai} = \log \frac{\pi_{Ai}}{(1 - \pi_{Ai})} \qquad i \in R \cup S$$

$$s_{Bi} \sim Bin(n_{Bi}, \pi_{Bi}), \qquad \phi_{Bi} = \log \frac{\pi_{Bi}}{(1 - \pi_{Bi})} \qquad i \in R \cup T$$

$$s_{Ci} \sim Bin(n_{Ci}, \pi_{Ci}), \qquad \phi_{Ci} = \log \frac{\pi_{Ci}}{(1 - \pi_{Ci})} \qquad i \in S \cup T$$

$$\phi_{Ai} - \phi_{Bi} = \theta_{ABi} \qquad i \in R$$

$$\phi_{Ai} - \phi_{Ci} = \theta_{ACi} \qquad i \in S$$

$$\phi_{Bi} - \phi_{Ci} = \theta_{BCi} \qquad i \in T$$

$$\theta_{ABi} \sim N(\mu_{AB}, \tau^2) \quad i \in R \tag{13}$$

$$\theta_{ACi} \sim N(\mu_{AC}, \tau^2) \quad i \in S \tag{14}$$

$$\theta_{BCi} \sim N(\mu_{BC}, \tau^2) \quad i \in T \tag{15}$$

$$\mu_{BC} = \mu_{AC} - \mu_{AB}$$

R, S and T are the three sets of trials comparing treatments A and B, A and C, and B and C, respectively.

Full conditionals may be determined for each parameter, and hence the Gibbs sampler used to make inferences about all three treatment differences and the common heterogeneity.

5.1. Trials comparing all three treatments

If a trial compares all three treatments A, B and C, then any assumptions about heterogeneity have implications on the relative efficacy parameters. Consider trial i, and the assumption of

equal heterogeneity parameters for each relative efficacy. The marginal distributions $\theta_{ABi} \sim N(\mu_{AB}, \tau^2)$, $\theta_{ACi} \sim N(\mu_{AC}, \tau^2)$ and $\theta_{BCi} \sim N(\mu_{BC}, \tau^2)$ together with the equality $\mu_{BC} = \mu_{AC} - \mu_{AB}$ imply that the covariance between any two θ_i's is equal to $\tau^2/2$. Thus, for such a trial (13)–(15) should be replaced with

$$\begin{pmatrix} \theta_{ABi} \\ \theta_{ACi} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_{AB} \\ \mu_{AC} \end{pmatrix}, \begin{pmatrix} \tau^2 & \tau^2/2 \\ \tau^2/2 & \tau^2 \end{pmatrix} \right)$$

and the distribution of θ_{BCi} is automatically fulfilled.

If the heterogeneity parameter cannot be assumed equal for each treatment comparison then the analysis is possible if some relationship between the heterogeneity parameters is specified. This might be in the form of the ratio of one to another. Otherwise, the assumptions $\theta_{ABi} \sim N(\mu_{AB}, \tau_1^2)$ and $\theta_{ACi} \sim N(\mu_{AC}, \tau_2^2)$ lead to a distribution for θ_{BCi} with variance $\tau_1^2 + \tau_2^2$.

5.2. Covariance between estimates for the general parametric approach

A trial with three treatment arms may be added to the exact binomial analysis by simply adding distributions for s_{Ci} and θ_{ACi}, a prior for μ_{AC} and the relation $\mu_{BC} = \mu_{AC} - \mu_{AB}$ to the specification of the model in BUGS. However, if the general parametric approach is used, then two estimates, say y_{ABi} and y_{ACi}, from one trial have data from treatment A in common and so are not statistically independent. Estimates of their covariance, denoted by D_{yi}, are presented in this section. To avoid unnecessary notation, all i subscripts are omitted, so relative efficacy estimates are denoted by y_{AB} and y_{AC}.

First, if maximum likelihood estimates are used, then, asymptotically,

$$D_{y,\mathrm{MLE}} = \frac{1}{s_A} + \frac{1}{n_A - s_A}$$

(see Bishop[33]).

If, on the other hand, the approximate maximum likelihood estimate based on the efficient score and Fisher's information is used, then an estimate of covariance derived here is appropriate. Retaining the same notational pattern, we have $y_{AB} = Z_{AB}/V_{AB}$ and $y_{AC} = Z_{AC}/V_{AC}$. Though we require an estimate, D_y, of the covariance between the relative efficacy estimates, in practice it is easier to derive an estimate, D_z, of the covariance between the efficient scores. Based on (8) it can be assumed that

$$\begin{pmatrix} Z_{AB} \\ Z_{AC} \end{pmatrix} \sim N\left(\begin{pmatrix} \theta_{AB} & V_{AB} \\ \theta_{AC} & V_{AC} \end{pmatrix}, \begin{pmatrix} V_{AB} & D_z \\ D_z & V_{AC} \end{pmatrix} \right).$$

It should be noted that $D_y = \mathrm{cov}(y_{AB}, y_{AC}) = V_{AB}^{-1} V_{AC}^{-1} D_z$, yielding an estimate for the desired covariance.

The two comparisons from the trial may be specified as the independent univariate distribution

$$y_{AB} \sim N(\theta_{AB}, v_{AB}) \tag{16}$$

and, from multivariate normal theory

$$y_{AC} \mid y_{AB} \sim N(\theta_{AC} + D_y v_{AB}^{-1}(y_{AB} - \theta_{AB}), \quad v_{AC} - D_y^2 v_{AB}^{-1}). \tag{17}$$

Applying this theory to binary data, we have, using (10),

$$D_z = \text{cov}\left(\frac{s_A n_B - n_A s_B}{n_A + n_B}, \frac{s_A n_C - n_A s_C}{n_A + n_C}\right)$$

$$= \frac{n_A n_B n_C}{(n_A + n_B)(n_A + n_C)} \pi_A(1 - \pi_A).$$

This follows from the assumed mutual independence of s_A, s_B and s_C. Estimating π_A by its maximum likelihood estimate s_A/n_A, an estimate of covariance between the two log odds ratio estimates is given by

$$D_y = \frac{s_A(n_A - s_A)n_B n_C}{V_{AB} V_{AC} n_A (n_A + n_B)(n_A + n_C)} \tag{18}$$

with the V's obtained using (11).

6. FIRST BLEEDING IN CIRRHOSIS

The methodology is illustrated using data from Pagliaro *et al.*[8] There are in total 26 trials investigating the use of beta-blockers and/or sclerotherapy in the prevention of first bleeding in cirrhosis. Nine of the trials make a comparison between beta-blockers and a control treatment; 19 between sclerotherapy and a control. Two trials compared all three treatments, and consequently only these two made a direct comparison between beta-blockers and sclerotherapy. Results from all the trials are given in Table I.

In the original paper the two pairwise comparisons involving the control treatment were investigated individually using the Peto fixed effect method. Beta-blockers and sclerotherapy were both found significantly superior to the control, the former more so than the latter. Significant heterogeneity was identified but a random effects analysis was not performed.

For the purposes of this example, the treatment difference of interest will be assumed to be that between beta-blockers and sclerotherapy. Although only two trials investigated this comparison, the scenario of having only two trials in a meta-analysis might be thought of as equivalent to performing the first interim analysis of a prospective cumulative meta-analysis. Estimates of the log odds ratio of bleeding on beta-blockers relative to sclerotherapy in these two trials are illustrated in Figures 3(*a*) and 4(*a*). The first graph gives approximate maximum likelihood estimates using (9); the second gives maximum likelihood estimates using (7), together with 95 per cent confidence intervals based on their estimated variances. Individual trial estimates are, for the Z/V method, $-1{\cdot}49$ (SE 0·65) for study 1 and $-0{\cdot}01$ (SE 0·44) for study 2, and for the MLE method, $-1{\cdot}72$ (SE 0·82) for study 1 and $-0{\cdot}01$ (SE 0·44) for study 2. In the first trial the approximate method gives a typically conservative estimate due to the underestimation of large log odds ratios by this method.

In the rest of this section a number of Bayesian meta-analyses are performed on these two trials, incorporating various combinations of the two sources of external information. Each is repeated using the general parametric approach with approximate maximum likelihood estimates of log odds ratios (Figure 3) and using the exact binomial approach (Figure 4). Summary statistics from the posterior distributions of μ and τ^2 appear in Table II.

6.1. Combining two trials of beta-blockers and sclerotherapy

Despite the obvious objections to performing a traditional meta-analysis on only two trials, their results are illustrated with broken lines in Figures 3(*b*) and 4(*b*). These will serve to demonstrate

Table I. Randomized trials of non-surgical treatment of first bleeding in cirrhosis: raw data and log odds ratios from Pagliaro *et al.*[8]

Study number	Number of patients			Log odds ratio estimates	
	Beta-blockers (A) bled/total	Sclerotherapy (B) bled/total	Control (C) bled/total	(Efficient score and Fisher's information) Comparison	Estimate (SE)
1	2/43	9/42	13/41	A–B	−1·49 (0·65)
				A–C	−1·84 (0·57)
2	12/68	13/73	13/72	A–B	−0·01 (0·44)
				A–C	−0·03 (0·44)
3	4/20	–	4/16	A–C	−0·29 (0·81)
4	20/116	–	30/111	A–C	−0·57 (0·32)
5	1/30	–	11/49	A–C	−1·48 (0·65)
6	7/53	–	10/53	A–C	−0·42 (0·53)
7	18/85	–	31/89	A–C	−0·68 (0·34)
8	2/51	–	11/51	A–C	−1·59 (0·59)
9	8/23	–	2/25	A–C	1·62 (0·71)
10	–	4/18	0/19	B–C	2·30 (1·06)
11	–	3/35	22/36	B–C	−2·30 (0·50)
12	–	5/56	30/53	B–C	−2·19 (0·41)
13	–	5/16	6/18	B–C	−0·10 (0·73)
14	–	3/23	9/22	B–C	−1·42 (0·67)
15	–	11/49	31/46	B–C	−1·82 (0·41)
16	–	19/53	9/60	B–C	1·12 (0·44)
17	–	17/53	26/60	B–C	−0·48 (0·39)
18	–	10/71	29/69	B–C	−1·39 (0·38)
19	–	12/41	14/41	B–C	−0·23 (0·47)
20	–	0/21	3/20	B–C	−2·21 (1·20)
21	–	13/33	14/35	B–C	−0·03 (0·50)
22	–	31/143	23/138	B–C	0·32 (0·30)
23	–	20/55	19/51	B–C	−0·04 (0·40)
24	–	3/13	12/16	B–C	−2·08 (0·75)
25	–	3/21	5/28	B–C	−0·26 (0·78)
26	–	6/22	2/24	B–C	1·32 (0·78)

what effects the incorporation of both types of external information can have on a meta-analysis of a small number of trials. The model is implemented in a Bayesian manner using the Gibbs sampler with a non-informative $IG(10^{-3}, 10^{-3})$ prior distribution for τ^2. This does not give a satisfactorily convergent chain, even after many iterations. Both approaches lead to erratic traces of simulated samples from the posterior distributions.

It is clearly inadvisable to attempt to combine only two studies in this manner. The main reason for the failure of the Gibbs sampler is the almost total lack of information regarding heterogeneity between the trials. The 95 per cent posterior probability intervals for τ^2 in both cases range from order 10^{-3} to order 10^2. If a meta-analysis of such a small number of studies is to be performed then imposing an influential prior distribution on τ^2 can provide enough robustness for the Gibbs sampler to produce acceptable results.

6.2. Analysis using empirical prior distribution for τ^2

Meta-analyses of the two beta-blocker versus sclerotherapy trials were repeated using the $IG(1{\cdot}0, 0{\cdot}35)$ prior distribution from studies in gastroenterology derived using the random effects

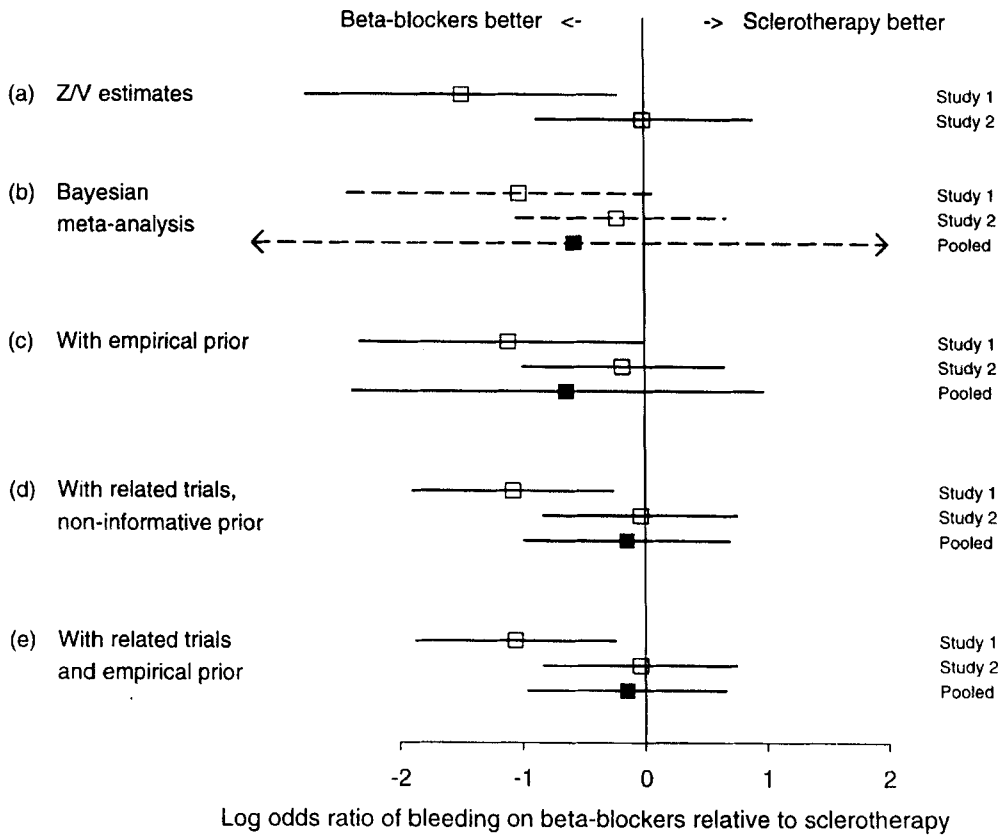

Figure 3. Meta-analyses of Pagliaro *et al.*[8] data using general parametric approach (Z/V method). Broken lines indicate lack of convergence of the Gibbs sampler. (*a*) Individual estimates with 95 per cent confidence intervals. (*b*) Bayesian meta-analysis using non-informative priors with 95 per cent posterior probability intervals. (*c*) Bayesian analysis with empirical prior for τ^2. (*d*) Bayesian analysis incorporating extra additional treatment comparisons using non-informative priors. (*e*) Bayesian analysis incorporating additional treatment comparisons using empirical prior for τ^2. All Bayesian results are based on a run of 15,000 iterations of the Gibbs sampler, following a burn-in of 1000

meta-analysis of meta-analyses approach described in Section 4.1. Results are given in Table II and illustrated in Figures 3(*c*) and 4(*c*). A non-informative $N(0, 10^3)$ prior distribution is again used for the relative efficacy parameter. In both approaches the convergence diagnostic tests were passed, and 'nice' sample traces from posterior distributions were produced.

The two approaches gave similar results, but for the slight underestimation of the treatment difference by the Z/V method, and, consequently, a slightly smaller estimate of the heterogeneity parameter. The posterior distributions of the individual trial relative efficacies are shrunk relative to the initial log odds ratio estimates and shrunk towards the overall mean. The shrinkage of the second study is less than that of the first due to it having more patients and hence providing more information.

6.3. Incorporating external trials

When the data were analysed in the original paper, studies 1 and 2 in Table I were pooled with the beta-blocker versus control trials and separately with the sclerotherapy versus control trials. This

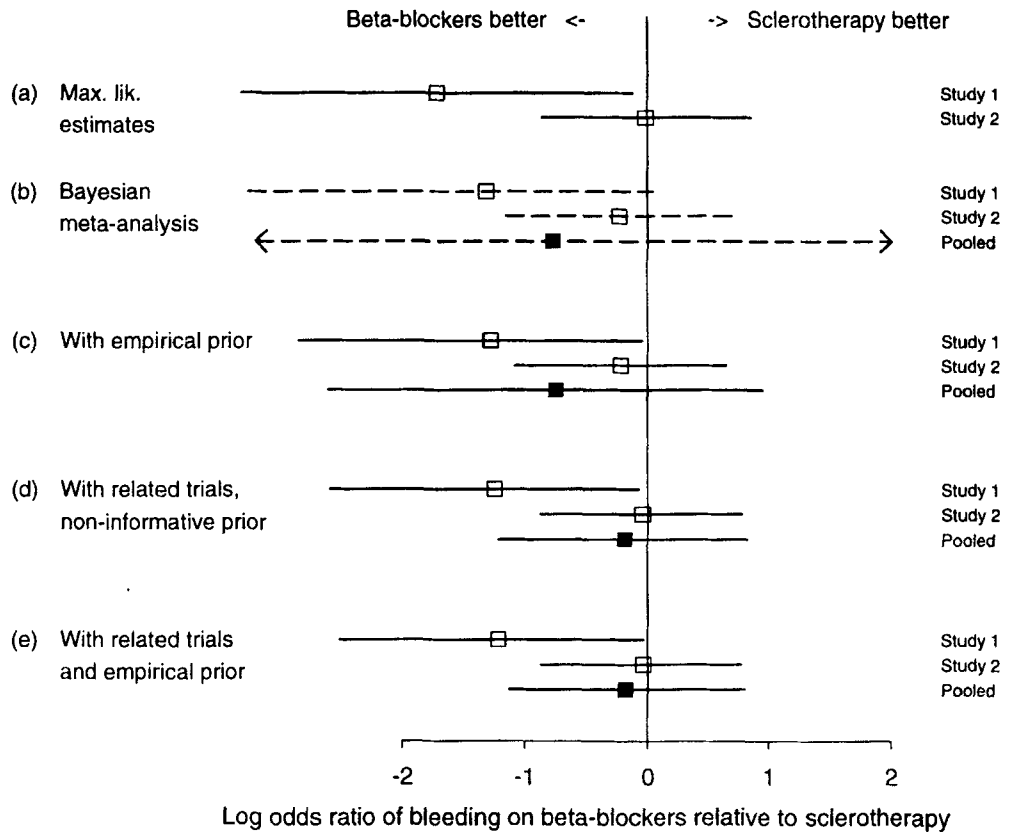

Figure 4. Maximum likelihood estimates of log odds ratio of bleeding on beta-blockers relative to sclerotherapy and meta-analyses of Pagliaro *et al.*[8] data using exact binomial approach. Broken lines indicate lack of convergence of the Gibbs sampler. (*a*) Maximum likelihood estimates of log odds ratios with 95 per cent confidence intervals. (*b*) Bayesian meta-analysis using non-informative priors with 95 per cent posterior probability intervals. (*c*) Bayesian analysis with empirical prior for τ^2. (*d*) Bayesian analysis incorporating extra additional treatment comparisons using non-informative priors. (*e*) Bayesian analysis incorporating additional treatment comparisons using empirical prior for τ^2. All Bayesian results are based on a run of 15,000 iterations of the Gibbs sampler, following a burn-in of 1000

implies the authors considered all trials in the data set 'poolable'. While this assumption has not been investigated by the current authors it is exploited here by incorporating all 26 trials into one analysis, assuming the same degree of heterogeneity for the three relative efficacy parameters beta-blockers versus sclerotherapy, beta-blockers versus control and sclerotherapy versus control.

The assumption of equal heterogeneity may be investigated by considering separate analyses of the pairwise comparisons of each active treatment with control. Using the exact binary method, Bayesian analyses using non-informative priors throughout give 95 per cent posterior probability intervals of (0·005, 4·6) for beta-blockers versus control and (0·61, 4·0) for sclerotherapy versus control. The latter clearly lies within the former, so we have no evidence that the true heterogeneity parameters should be different. Analysing the two sets using the general parametric approach yields similar intervals, though the latter extends beyond the former at the upper end by just over 0·4, which is not enough to raise particular concern.

Table II. Results of meta-analyses of data in Table I. Posterior mean log odds ratio, μ, and posterior median heterogeneity parameter, τ^2, with 95 per cent posterior probability intervals. Non-informative prior is IG(10^{-3}, 10^{-3}); empirical prior distribution is IG(1·0, 0·35). All results based on 15,000 iterations of the Gibbs sampler following a burn-in period of 1000 iterations

Trials included (n)	Prior distribution	Estimation method	Mean of μ 95 per cent probability interval	Median of τ^2 95 per cent probability interval
Studies 1 and 2 only (2)	non-informative	Z/V	−0·58 (−5·62, 5·03)	0·43 (0·00, 321)
		exact	−0·77 (−7·74, 6·23)	0·88 (0·00, 625)
	empirical	Z/V	−0·65 (−2·39, 0·97)	0·52 (0·10, 6·92)
		exact	−0·74 (−2·61, 0·95)	0·57 (0·10, 7.43)
With extra comparisons (26)	non-informative	Z/V	−0·15 (−0·99, 0·70)	0·91 (0·42, 2·00)
		exact	−0·18 (−1·22, 0·82)	1·33 (0·59, 3·00)
	empirical	Z/V	−0·15 (−0·97, 0·66)	0·84 (0·39, 1·78)
		exact	−0·18 (−1·13, 0·79)	1·20 (0·54, 2·64)

As an aside, the differences between the posterior means of the comparisons with control, to give crude estimates of the relative efficacy of beta-blockers and sclerotherapy, are −0·07 for the general parametric approach and −0·06 for the exact binomial approach. These give a rough guide to the likely magnitude of treatment difference which would follow incorporation of the 24 additional trials. The relative sizes of the studies and an assumption of equal heterogeneity parameters also affect the estimation of relative efficacy.

Incorporating the additional trials into either approach is straightforward. Identifying treatments A, B and C of Section 5 with beta-blockers, sclerotherapy and control, respectively, the relative efficacy, μ, of interest may be represented by μ_{AB}. The ith trial investigates either θ_{ACi}, θ_{BCi} or, in the case of the two three-armed studies, both θ_{ABi} and θ_{ACi}. In the exact binomial approach each of the three θ's is modelled. In the general parametric approach, equations (16) and (17) are used, with covariances estimated using equation (18).

The results of adding all extra comparisons of active treatment versus control to the model, and returning to the non-informative IG(10^{-3}, 10^{-3}) prior distribution for τ^2, are presented in Table II and Figures 3(d) and 4(d). Immediately obvious is that the posterior overall mean log odds ratio is much nearer to zero in both cases. An explanation for this lies partly in the results of the pairwise analyses for the active treatments separately. It would appear that study 1 produced an unusually positive result in favour of beta-blockers. A second observation is that the posterior distributions for the heterogeneity parameter are significantly tighter. This shrinks the estimates toward their mean, and, coupled with extra, indirect, information concerning the relative efficacy of beta-blockers and sclerotherapy, tightens the posterior distributions of all parameters.

6.4. Analysis of example data set with external trials and an empirical prior for τ^2

Finally, the empirical prior distribution IG(1·0, 0·35) is imposed on τ^2 in addition to incorporating the external trials. Results are given in Table II and Figures 3(e) and 4(e). The posterior distribution of τ^2 is tighter still, and shifted towards zero, suggesting that more heterogeneity was observed in the current set of trials than would normally be expected. The prior has little effect on the posterior estimates of relative efficacy, but leads to smaller standard errors.

7. DISCUSSION

Two sources of external information relevant to the meta-analysis of a set of randomized trials have been described, and illustrated in the context of an example in gastroenterology. The paper has discussed their incorporation into two approaches of estimating log odds ratios and heterogeneity for binary data, namely a general parametric approach requiring only two summary statistics from each trial and an exact binomial approach requiring individual patient data. The second is clearly more suitable when the data are available, since it avoids the assumption of known variances and covariances. However, there are a number of reasons why the first is included. First, it may be used when some published trial papers do not include a full breakdown of the results; second, the approach is applicable to many other types of data; and third, there may be occasions when there are too few data for the Gibbs sampler for the explicit binomial approach to converge, but where it does for the general parametric approach. However, if patient numbers and success rates are not available then covariances can often not be calculated, and three treatment trials cannot be properly included in the parametric approach.

The approximate maximum likelihood estimation method (Z/V) is used in the example because the Peto $O - E$ statistic is most commonly used, and it is interesting to compare its performance with a more data-specific approach. The example illustrates the underestimation of treatment differences from which the estimate suffers.

The use of historical meta-analyses to form a prior distribution for the heterogeneity involves the assumption that heterogeneity between similar sets of trials is likely to be similar. This assumption needs careful consideration, but in most cases may be reasonable. Given this assumption the prior is entirely data based. It is particularly relevant when there are few trials in the meta-analysis data set, in which case it is difficult to obtain a good estimate of the heterogeneity, and thus difficult to account for it realistically in the pooling of trial results.

The various approaches to formulating a prior distribution have different merits. The *ad hoc* method of finding a close fit to the empirical cumulative distribution function by trial and error is simple and in the case of the gastroenterology example yielded a similar prior to the more complicated methods. It may therefore not be worth the extra effort of performing a meta-analysis of meta-analyses. Of the three parametric methods, the distribution of τ^2_{new} assuming random effects for the individual trial heterogeneity parameters would be most appropriate, even though this also requires approximation of the kernel density with a parametric distribution. Using the posterior medians of the parameters α and λ is unwise given the wide probability intervals surrounding them. The fixed effect model suffers from the same drawback as a traditional fixed effect meta-analysis when heterogeneity is present. In our example the distribution this method produced was not suitably diffuse.

It would, of course, be possible to incorporate the information about heterogeneity from historical meta-analyses into the point estimate of τ^2 in the frequentist meta-analysis of Section 2.1. For example, the median of an empirical prior distribution might be used. However, this takes no account of the uncertainty surrounding the estimate. Also, there is no prescribed procedure for the obviously desirable combination of such a statistic with an estimate of the heterogeneity in the present data set.

Including related trials in the meta-analysis provides more information concerning the treatment difference of interest. In some cases this may produce a more precise estimate, in others it may widen posterior probability intervals. In our meta-analysis of the Pagliaro data, we found the former. If the latter occurs this suggests that either the trials making the direct comparison or the additional trials are in some sense unusual, or that the two sets of studies are fundamentally different. In practice this would lead to further investigation of the circumstances of the trials.

Many meta-analysis papers include data from three or more treatments, but only consider pairwise comparisons of, say treatment A with control and treatment B with control. There would seem to be little reason not to combine all treatments into one analysis. If equal heterogeneity parameters can be assumed then the methods presented in this paper may be used. However, when there are few trials in a meta-analysis there is little evidence to be able to accept or reject this assumption. We may not be particularly concerned if this assumption does not appear valid because we are not interested in estimating the heterogeneity parameter, but in using it as a tool to give more conservative posterior probability intervals for the relative efficacy parameters.

If an assumption of equal heterogeneity parameters does not appear to be reasonable, then the additional trials can still be used. It may be possible to specify a ratio of one heterogeneity parameter to another. Alternatively, the information can be used to formulate a data based prior distribution for the overall relative efficacy parameter, μ. Just as results from external trials may be formed into a prior distribution, results from historical meta-analyses may be incorporated directly into the model in a similar manner to the meta-analysis of meta-analyses used to infer about τ^2.

All analyses presented in the paper are based on data and not on subjective opinion, and indeed a non-informative prior distribution has always been placed on treatment difference parameters. The methods are generalizable to more than three treatments, to other data types and to other situations in which external information is available. The ease of use and availability of the BUGS software makes it a possibility for any meta-analysis to include as much relevant information as possible.

ACKNOWLEDGEMENT

The first author was supported by a Medical Research Council research studentship.

REFERENCES

1. DerSimonian, R. and Laird, N. 'Meta-analysis in clinical trials', *Controlled Clinical Trials*, **7**, 177–188 (1986).
2. Whitehead, A. and Whitehead, J. 'A general parametric approach to the meta-analysis of randomized clinical trials', *Statistics in Medicine*, **10**, 1665–1677 (1991).
3. Collett, D. *Modelling Binary Data*, Chapman and Hall, London, 1991.
4. Carlin, J. B. 'Meta-analysis for 2×2 tables: a Bayesian approach', *Statistics in Medicine*, **11**, 141–158 (1992).
5. *BUGS Manual 0.30*, MRC Biostatistics Unit, Cambridge, 1994.
6. Begg, C. B. and Pilote, L. 'A model for incorporating historical controls into a meta-analysis', *Biometrics*, **47**, 899–906 (1991).
7. Smith, T. C. 'Interpreting evidence from multiple randomised and non-randomised studies', Ph.D. thesis, University of Cambridge, 1995.
8. Pagliaro, L., D'Amico, G., Sorensen, T. I., Lebrec, D., Burroughs, A. K., Morabito, A., Tine, F., Politi, F. and Traina, M. 'Prevention of first bleeding in cirrhosis. A meta-analysis of randomized trials of nonsurgical treatment', *Annals of Internal Medicine*, **117**, 59–70 (1992).
9. Whitehead, J. *The Design and Analysis of Sequential Clinical Trials*, 2nd edn, Ellis Horwood, Chichester, 1992.
10. Yusuf, S., Peto, R., Lewis, J., Collins, R. and Sleight, T. 'Beta blockade during and after myocardial infarction: an overview of the evidence', *Progress in Cardiovascular Disease*, **27**, 335–371 (1985).
11. Hardy, R. J. and Thompson, S. G. 'A likelihood approach to meta-analysis with random effects', *Statistics in Medicine*, **15**, 619–629 (1996).
12. *EGRET Reference Manual: First Draft*, Statistics and Epidemiology Research Corporation, Seattle, 1990.
13. Smith, T. C., Speigelhalter, D. J. and Thomas, A. 'Bayesian approaches to random effects meta-analysis: a comparative study', *Statistics in Medicine*, **14**, 2685–2699 (1995).

14. Box, G. and Tiao, G. *Bayesian Inference in Statistical Analysis*, Addison-Wesley, Reading, MA, 1973.
15. DuMouchel, W. and Waternaux, C. 'Hierarchical models for combining information and for meta-analyses (discussion)', in Bernado, J. M. *et al.*(eds.), *Bayesian Statistics 4*, Clarendon Press, Oxford, 1992, pp. 338–341.
16. Gelfand, A. E. and Smith, A. F. M. 'Sampling based approaches to calculating marginal densities', *Journal of the American Statistical Association*, **85**, 398–409 (1990).
17. Thomas, A., Spiegelhalter, D. and Gilks, W. '*BUGS*: A program to perform Bayesian inference using Gibbs sampling', in Bernado, J. M. *et al.* (eds.), *Bayesian Statistics 4*, Clarendon Press, Oxford, 1992, pp. 837–842.
18. Raftery, A. E. and Lewis, S. M. 'How many iterations of the Gibbs sampler?', in Bernado, J. M. *et al.*(eds.), *Bayesian Statistics 4*, Clarendon Press, Oxford, 1992, pp. 763–773.
19. Geweke, J. 'Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments', in Bernado, J. M. *et al.* (eds.), *Bayesian Statistics 4*, Clarendon Press, Oxford, 1992, pp. 169–193.
20. Heidelberger, P. and Welch, P. D. 'Simulation run length control in the presence of an initial transient', *Operations Research*, **31**, 1109–1144 (1983).
21. Best, N. G., Cowles, M. K. and Vines, S. K. *CODA Manual version 0.30*, MRC Biostatistics Unit, Cambridge, 1995.
22. Henry, D. A. and O'Connell, D. L. 'Effects of fibrinolytic inhibitors on mortality from upper gastrointestinal haemorrhage', *British Medical Journal*, **298**, 1142–1146 (1989).
23. Lacroix, J., Infante-Rivard, C., Jenicek, M. and Gauthier, M. 'Prophylaxis of upper gastrointestinal bleeding in intensive care units: A meta-analysis', *Critical Care Medicine*, **17**, 862–869 (1989).
24. Sacks, H. S., Chalmers, T. C., Blum, A. L., Berrier, J. and Pagano, D. 'Endoscopic haemostasis. An effective therapy for bleeding peptic ulcers', *Journal of the American Medical Association*, **264**, 494–499 (1990).
25. Ricca-Rosellini, S. and Miglio, F. 'Beta-blockers for the prevention of variceal haemorrhage in patients with cirrhosis: an updated meta-analysis of randomized controlled trials', *Italian Journal of Gastroenterology*, **23**, 408–415 (1991).
26. Cook, D. J., Witt, L. G., Cook, R. J. and Guyatt, G. H. 'Stress ulcer prophylaxis in the critically ill: a meta-analysis', *American Journal of Medicine*, **91**, 519–527 (1991).
27. Poynard, T., Cales, P., Pasta, L., Ideo, G., Pascal, J. P., Pagliaro, L. and Lebrec, D. 'Beta-adrenergic-antagonist drugs in the prevention of gastrointestinal bleeding in patients with cirrhosis and esophageal varices. An analysis of data and prognostic factors in 589 patients from four randomized clinical trials. Franco-Italian Multicenter Study Group', *New England Journal of Medicine*, **324**, 1532–1538 (1991).
28. Spina, G. P., Henderson, J. M., Rikkers, L. F., Teres, J., Burroughs, A. K., Conn, H. O., Pagliaro, L. and Santambrogio, R. 'Distal spleno-renal shunt versus endoscopic sclerotherapy in the prevention of variceal rebleeding. A meta-analysis of 4 randomized clinical trials', *Journal of Hepatology*, **16**, 338–345 (1992).
29. Van Ruiswyk, J. and Byrd, J. C. 'Efficacy of prophylactic sclerotherapy for prevention of a first variceal hemorrhage', *Gastroenterology*, **102**, 587–597 (1992).
30. Chalmers, T. C. and Lau, J. 'Meta-analytic stimulus for changes in clinical trials', *Statistical Methods in Medical Research*, **2**, 161–172 (1993).
31. Lam, S. K. and Lai, K. C. 'Endoscopic haemostasis for gastrointestinal bleeding: the dawning of a new era', *Journal of Gastroenterology and Hepatology*, **9**, 69–74 (1994).
32. DuMouchel, W. H. 'Bayesian meta-analysis', in Berry, D. A. (ed.), *Statistical Methodology in the Pharmaceutical Sciences*, Marcel Dekker, New York, 1990, pp. 509–529.
33. Bishop, Y. M. M., Fienberg, S. E. and Holland, P. W. *Discrete Multivariate Analysis Theory and Practice*, M.I.T. Press, Cambridge, Mass., 1975.