STATISTICS IN MEDICINE, VOL. 16, 753–768 (1997)

INCORPORATING VARIABILITY IN ESTIMATES OF HETEROGENEITY IN THE RANDOM EFFECTS MODEL IN META-ANALYSIS

B. J. BIGGERSTAFF

Statistical Engineering Division, National Institute of Standards and Technology, Boulder, CO 80303-3328, U.S.A.

AND

R. L. TWEEDIE

Department of Statistics, Colorado State University, Fort Collins, Colorado 80523, U.S.A.

SUMMARY

When combining results from separate investigations in a meta-analysis, random effects methods enable the modelling of differences between studies by incorporating a heterogeneity parameter τ^2 that accounts explicitly for across-study variation. We develop a simple form for the variance of Cochran's homogeneity statistic Q, leading to interval estimation of τ^2 utilizing an approximating distribution for Q; this enables us to extend the point estimation of DerSimonian and Laird. We also develop asymptotic likelihood methods and compared them with this method. We then use these approximating distributions to give a new method of calculating the weight given to the individual studies' results when estimating the overall mean which takes into account variation in these point estimates of τ^2. Two examples illustrate the methods presented, where we show that the new weighting scheme is between the standard fixed and random effects models in down-weighting the results of large studies and up-weighting those of small studies. © 1997 by John Wiley & Sons, Ltd.

1. INTRODUCTION

Utilization of meta-analytic techniques to combine results and information from separate quantitative investigations has become increasingly common in the past 10 years.[1–4] There are, however, well-founded concerns regarding the application of these techniques to sets of studies that, when taken on the whole, appear discordant in their individual conclusions.[5,6]

The random effects (RE) model in meta-analysis, which we study here, has been suggested as a way to model known differences between studies,[3,7] such as study design, different within-study matching protocols, different treatment protocols, or perhaps even gender or culture differences between study participants. It does this by incorporating a variance parameter τ^2 explicitly to account for across-study variation. Adoption of the RE model in meta-analysis permits extension of inferences to a broader population of studies than does the so-called 'fixed effects' (FE) model[3,4,7] which excludes the parameter τ^2 from the model; for this reason, the application of the FE model in meta-analytic contexts has been called into question.[3]

In this paper we investigate methods of interval estimation for τ^2 and methods of assessing the sensitivity of the RE model to the underlying variation in estimates of τ^2 when estimating the

CCC 0277–6715/97/070753–16$17.50

Received February 1995
Revised May 1996

overall mean parameter μ. Point estimation of τ^2 has been investigated systematically by DerSimonian and Laird in reference 7, and here we build on this work.

We assume that we have collected K studies, from which we have extracted summary information concerning inference on a common effect of interest; in the examples we use as illustration, this is the true relative risk of pre-eclampsia when using diuretics,[8] or the true relative risk of lung cancer as a result of exposure to environmental tobacco smoke.[9] We define the RE model as follows: for $i = 1, \ldots, K$, let

$$Y_i = \mu_i + \varepsilon_i, \quad \varepsilon_i \sim \mathrm{N}(0, \sigma_i^2)$$

$$\mu_i = \mu + E_i, \quad E_i \sim \mathrm{N}(0, \tau^2) \tag{1}$$

where the Y_is are the relevant effects measures from the individual studies (for example, logarithms of observed odds ratios or mean differences, see reference 2); the σ_i^2s are the variances of these individual measures; μ_i is the true mean in study i; μ is the true underlying mean (generally the parameter of interest); τ^2 is the between- or across-study variance; and the ε_i and E_i are mutually independent, normally distributed error terms. In the frequentist setting, to which we confine our discussion, it is common practice to assume that the individual studies provide good estimates of the within-study variance, so we assume that the σ_i^2 are known, and then we simply replace these with the reported estimates from the studies.

As a method of assessing the hypothesis that the heterogeneity parameter τ^2 is actually zero, Cochran[10] proposed a now widely-used test of homogeneity based on the statistic

$$Q = \sum w_i (Y_i - \hat{\mu})^2 \tag{2}$$

where $w_i = \sigma_i^{-2}$, and $\hat{\mu} = (\sum w_i Y_i)/\sum w_i$ is the weighted least squares estimator of μ ignoring the variation in the μ_i. Under the hypothesis of homogeneity ($\tau^2 = 0$) Q follows a χ^2_{K-1} distribution;[10] a large observed value of the statistic Q relative to this distribution indicates rejection of the hypothesis of homogeneity, which should indicate the appropriateness of the RE model. (The same statistic Q gives a test of homogeneity for the FE model, though the null hypothesis in this case is that all the μ_i are equal.)

It has been noted,[5] however, that this test is low in power, so that one should not interpret the failure to reject the hypothesis of homogeneity as strong indication of homogeneity. The DerSimonian–Laird point estimate $\hat{\tau}^2_{\mathrm{DL}}$ of τ^2 given in equation (6) is, however, unaffected by the low power of this homogeneity test, as one may find $\hat{\tau}^2_{\mathrm{DL}} > 0$ even if the result of the test is non-significance.

In Section 2.1 we give a simple form for the variance of Q. This variance, which yields immediately the variance of $\hat{\tau}^2_{\mathrm{M}}$, an unbiased moment estimate for τ^2 we define below, seems to be new to meta-analytic applications; Raudenbush comments that the standard error for the estimate of τ^2 is 'not available in the method of moments case' (reference 11, p. 319).

We then derive an approximate distribution for $\hat{\tau}^2_{\mathrm{M}}$, which is the non-truncated version of $\hat{\tau}^2_{\mathrm{DL}}$,[7] using an approximating gamma distribution for Q, made possible with the availability of $\mathrm{var}(Q)$. This gives an approximate CI for τ^2, computed using this approximate distribution for $\hat{\tau}^2_{\mathrm{M}}$.

In Section 2.2 we derive two alternative CIs for τ^2 with likelihood methods, one using an asymptotic result for the generalized likelihood ratio statistic Λ_{K} and the other asymptotic distribution of the maximum likelihood estimator $\hat{\tau}^2_{\mathrm{ML}}$ for τ^2. We also give the asymptotic likelihood ratio test of homogeneity. These enable us to make further inferences about the hypothesis $\tau^2 = 0$ and additionally about how large τ^2 is apt to be.

We can then consider how the observed variability in estimates of τ^2 affect inference about the underlying value μ. It is standard[2,3,7] when estimating μ to use a point estimate for τ^2 in

$$\hat{\mu}(\tau^2) = \frac{\sum(\sigma_i^2 + \tau^2)^{-1} Y_i}{\sum(\sigma_i^2 + \tau^2)^{-1}} \tag{3}$$

and

$$\text{var}[\hat{\mu}(\tau^2)] = \frac{1}{\sum(\sigma_i^2 + \tau^2)^{-1}} \tag{4}$$

to construct point estimates and then confidence intervals (CI) for μ using normal distribution theory. In this RE approach, sensitivity analysis beyond performing Cochran's test of homogeneity has not been routinely performed, in comparison to the common practice in Bayesian modelling.[3,12,13] Our approach, developed in Section 3, estimates the weights $(\sigma_i^2 + \tau^2)^{-1}$ in equations (3) and (4) taking into account inherent variability of the point estimators of τ^2 used there.

Finally, in Section 4 we give two examples. We illustrate computations of the CIs, and we compare the standard DerSimonian–Laird method with the proposed weighting scheme for estimating μ. The first example uses data from a meta-analysis of the studies of prevention of pre-eclampsia with diuretics taken from Thompson and Pocock.[5] The second example uses data from a meta-analysis in Mengersen *et al.*[9] of the studies of the relative risk of lung cancer associated with exposure to environmental tobacco smoke (ETS) in the home.

Appendix II provides code using SAS[14] statistical software as a guide to implementation of the more difficult computations.

There are many other issues to take into account when performing a meta-analysis, such as identifying and accounting for publication bias or for specific biases within individual studies in a particular application. While we acknowledge these general shortcomings, the proposed analysis gives the meta-analyst a more formal approach to evaluation of the role heterogeneity plays in application of the RE model.

2. ESTIMATION OF τ^2

2.1. Method of moments approach for estimating τ^2

DerSimonian and Laird[7] developed a method of moments estimator $\hat{\tau}^2_{\text{DL}}$ for τ^2 using Q in equation (2) which is easily calculable and has been highly recommended.[3] In constructing CIs on μ with this estimator, Larholt *et al.*[15] utilize Satterthwaite's approximation[16] for the distribution of Q after computing var(Q), but only under the restriction that each $\sigma_i^2 \ll \tau^2$. We have found this restriction inappropriate in applications we have encountered, however (see Section 4), and the exact form of var(Q) that we now derive needs no such assumption.

We can easily find (reference 17, equations (5–6)) that the expected value of Q is

$$\mathbb{E}(Q) = (K-1) + \left(S_1 - \frac{S_2}{S_1}\right)\tau^2 \tag{5}$$

where $S_r = \sum w_i^r$. Note that $\mathbb{E}(Q) \geqslant 0$, since the term $S_1 - (S_2/S_1) \geqslant 0$, and that equation (5) is true in the more general case that the errors ε_i and E_i do not follow a normal distribution. From equation (5) we can readily derive the DerSimonian–Laird estimator $\hat{\tau}^2_{\text{DL}}$[7] as

$$\hat{\tau}^2_{\text{DL}} = \max\left\{0, \frac{Q-(K-1)}{S_1 - \left(\frac{S_2}{S_1}\right)}\right\}. \tag{6}$$

Let $\hat{\tau}^2_{\mathrm{M}} = [Q - (K-1)]/[S_1 - (S_2/S_1)]$ denote the moment estimator on which the DerSimonian–Laird estimator is based; note $\hat{\tau}^2_{\mathrm{M}}$ may be negative, when it is naturally set to zero for point estimation as in equation (6). We use an approximating distribution for the non-truncated $\hat{\tau}^2_{\mathrm{M}}$ to construct confidence limits for τ^2 by calculating $\mathrm{var}(Q)$ and then fitting an approximate distribution by matching the first two moments.

Theorem 1 Under the model in equations (1), the variance of Cochran's Q given by equation (2) is

$$\mathrm{var}(Q) = 2(K-1) + 4\left(S_1 - \frac{S_2}{S_1}\right)\tau^2 + 2\left(S_2 - 2\frac{S_3}{S_1} + \frac{S_2^2}{S_1^2}\right)\tau^4 \tag{7}$$

with $S_r = \sum w_i^r$.

A proof appears in the Appendix. Rao *et al.* (reference 17, p. 92) also provide a general formula for this variance, which they prove using results on moments of quadratic forms. The proof given here is more direct, and the final version of $\mathrm{var}(Q)$ in equation (7) is substantially less cumbersome than that of Rao *et al.* From this variance one readily computes

$$\mathrm{var}(\hat{\tau}^2_{\mathrm{M}}) = \mathrm{var}(Q)/[S_1 - (S_2/S_1)]^2.$$

Larholt *et al.* (reference 15, Appendix) also compute a form for $\mathrm{var}(Q)$, assuming $\sigma_i^2 \ll \tau^2$ for $i = 1, \ldots, K$. Under this assumption they arrive at

$$\mathrm{var}(Q) \approx 2\left(S_2 - 2\frac{S_3}{S_1} + \frac{S_2^2}{S_1^2}\right)\tau^4 \tag{8}$$

clearly an approximation to the exact quantity in equation (7).

Given $\mathbb{E}(Q)$ and $\mathrm{var}(Q)$, we now approximate the distribution of Q with a gamma distribution with shape parameter r and scale parameter λ; this is essentially Satterthwaite's method.[16] Setting $\mathbb{E}(Q) = r/\lambda$ and $\mathrm{var}(Q) = r/\lambda^2$ and solving for r and λ, and emphasizing the dependence on τ^2, we have

$$r(\tau^2) = \frac{\{\mathbb{E}(Q)\}^2}{\mathrm{var}(Q)} \quad \text{and} \quad \lambda(\tau^2) = \frac{\mathbb{E}(Q)}{\mathrm{var}(Q)}.$$

Note that when $\tau^2 = 0$, $\mathbb{E}(Q) = K - 1$ and $\mathrm{var}(Q) = 2(K-1)$, so that we 'approximate' the distribution of Q with a gamma $(\frac{1}{2}K - \frac{1}{2}, \frac{1}{2})$, or χ^2_{K-1}, distribution, which is the true distribution of Q in this case.

Based on this approximate distribution for Q, an approximate distribution for $\hat{\tau}^2_{\mathrm{M}}$ is a location-shifted, scaled, gamma distribution. The probability density function $f_{\mathrm{M}}(\cdot\,; \tau^2)$ of $\hat{\tau}^2_{\mathrm{M}}$ under this distributional assumption is

$$f_{\mathrm{M}}(t; \tau^2) = c\,\frac{[\lambda(\tau^2)]^{r(\tau^2)}}{\Gamma(r(\tau^2))}[ct + K - 1]^{r(\tau^2)-1}\,\mathrm{e}^{-\lambda(\tau^2)[ct+K-1]}\,I_{\left[-\frac{K-1}{c}, \infty\right)}(t)$$

for parameter $\tau^2 \geqslant 0$, where $c = [S_1 - (S_2/S_1)]$ and $I_A(\cdot)$ is the indicator of the set A.

We can use this density function to construct a CI for τ^2 as follows. Define functions $L(\tau^2)$ and $U(\tau^2)$ by

$$\begin{aligned} L(\tau^2) &= \int_{\hat{\tau}_m^2}^{\infty} f_M(t;\tau^2)\,dt \\ &= \int_{\lambda(\tau^2)[c\hat{\tau}_m^2+K-1]}^{\infty} \frac{1}{\Gamma(r(\tau^2))}\,u^{r(\tau^2)-1}\,e^{-u}\,du \\ U(\tau^2) &= \int_{-\frac{K-1}{c}}^{\hat{\tau}_m^2} f_M(t;\tau^2)\,dt \\ &= \int_{0}^{\lambda(\tau^2)[c\hat{\tau}_m^2+K-1]} \frac{1}{\Gamma(r(\tau^2))}\,u^{r(\tau^2)-1}\,e^{-u}\,du \end{aligned}$$

where $\hat{\tau}_m^2$ is the observed value of the statistic $\hat{\tau}_M^2$.

A $100(1-\alpha_1-\alpha_2)$ per cent CI for τ^2 based on this approximate distribution is now

$$(\hat{\tau}_L^2 = \tau_l^2, \hat{\tau}_U^2 = \tau_u^2)$$

where τ_l^2 and τ_u^2 are solutions for τ^2 in the equations $L(\tau^2)=\alpha_1$ and $U(\tau^2)=\alpha_2$, respectively.[18] (We present two-sided CIs for τ^2 throughout. The methods given may be altered in the natural way to yield one-sided CIs for τ^2 when these are desired.)

Numerical routines to find these solutions are not difficult to code; mainstream statistical computing packages generally include the gamma distribution function (also called the incomplete gamma function), and complete routines in C appear in reference 19. In Appendix II we provide methods of performing these computations in SAS.[14]

2.2. Likelihood approach for estimating τ^2

Alternative ways of finding CIs follow from likelihood methods. The likelihood function for observations $\{y_1, \ldots, y_K\}$ from the model in equation (1) is

$$f(y_1, \ldots, y_K; \mu, \tau^2) = (2\pi)^{-\frac{K}{2}} \prod_{i=1}^{K} \left(\frac{1}{\sigma_i^2+\tau^2}\right)^{1/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{K}\frac{(y_i-\mu)^2}{\sigma_i^2+\tau^2}\right\}$$

recalling that we take the σ_i^2 as known. The generalized likelihood ratio statistic Λ_K for the point null hypothesis $H_0:\tau^2=\tau_0^2$ then takes the form

$$\Lambda_K(\tau_0^2) = \prod_{i=1}^{K}\left(\frac{\sigma_i^2+\hat{\tau}_{ML}^2}{\sigma_i^2+\tau_0^2}\right)^{1/2} \exp\left[-\frac{1}{2}\sum_{i=1}^{K}\left\{\frac{(y_i-\hat{\mu}_0)^2}{\sigma_i^2+\tau_0^2}-\frac{(y_i-\hat{\mu}_{ML})^2}{\sigma_i^2+\hat{\tau}_{ML}^2}\right\}\right]$$

where $\hat{\mu}_{ML}$ and $\hat{\tau}_{ML}^2$ are the maximum likelihood estimates (MLE) of μ and τ^2, respectively, and where

$$\hat{\mu}_0 = \frac{\sum(\sigma_i^2+\tau_0^2)^{-1}y_i}{\sum(\sigma_i^2+\tau_0^2)^{-1}}.$$

We must find the MLEs of both μ and τ^2 numerically, though this is neither difficult nor computationally expensive; an iterative scheme for doing so appears in reference 3.

We can also derive a test of homogeneity and a CI using Λ_K together with the fact that $\lambda_K = -2\log(\Lambda_K)$, under the homogeneity hypothesis $\tau^2=0$, is asymptotically distributed as χ_1^2.[18]

An asymptotic $100(1-\alpha)$ per cent CI for τ^2 is given by the set

$$C_{1-\alpha} = \{\tau^2 : \lambda_K(\tau^2) \leqslant \chi^2_{1;1-\alpha}\}$$

where $\chi^2_{1;1-\alpha}$ is the $100(1-\alpha)$th percentile point of a χ^2_1 distribution.[18]

We can find an alternative asymptotic CI for τ^2 by arguing that the asymptotic distribution of the MLE of τ^2 is normally distributed with mean τ^2 and variance equal to the inverse of the Fisher information. We can readily show[2,17] that this variance is $2/\sum(\sigma_i^2+\tau^2)^{-2}$, so that a $100(1-\alpha)$ per cent asymptotic symmetric CI for τ^2 is

$$(\hat{\tau}^2_{\mathrm{ML}} - z_{(1-\frac{\alpha}{2})}\sqrt{\{2/\textstyle\sum(\sigma_i^2+\hat{\tau}^2_{\mathrm{ML}})^{-2}\}},\ \hat{\tau}^2_{\mathrm{ML}} + z_{(1-\frac{\alpha}{2})}\sqrt{\{2/\textstyle\sum(\sigma_i^2+\hat{\tau}^2_{\mathrm{ML}})^{-2}\}})$$

where $z_{(1-\frac{\alpha}{2})}$ is the $100(1-\frac{\alpha}{2})$th percentile point of the normal distribution.

3. ESTIMATION OF μ

As noted in Section 1, when estimating μ one typically substitutes a point estimate of τ^2 for τ^2 in equations (3) and (4) and then applies normal theory to obtain CIs. This ignores variation in the estimate of τ^2, and we address this problem in this section.

To facilitate presentation, write the true weight $(\sigma_i^2+\tau^2)^{-1}$ associated with study i in equations (3) and (4) more concisely as $w_i(\tau^2)$. The w_i of Section 2.1 and Theorem 1 are therefore simply $w_i(0)$ in this augmented notation, and this should cause no confusion in the derivations below. Also, we use only $\hat{\tau}^2_{\mathrm{DL}}$ in this section, though we could apply similar methods using the asymptotic normal distribution for $\hat{\tau}^2_{\mathrm{ML}}$ as discussed in Section 2.2.

The usual estimate of the true weight $w_i(\tau^2)$ is $w_i(\hat{\tau}^2_{\mathrm{DL}})$.[3,7] To utilize the distributional properties derived for $\hat{\tau}^2_{\mathrm{DL}}$ above in an effort to account for variation in $\hat{\tau}^2_{\mathrm{DL}}$, we propose a different set of weights $w_i^*(\tau^2)$ derived as follows.

With the approximating density $f_{\mathrm{M}}(\cdot;\tau^2)$ for $\hat{\tau}^2_{\mathrm{M}}$, one can write a 'density' form $f_{\mathrm{DL}}(t;\tau^2)$ for $\hat{\tau}^2_{\mathrm{DL}}$ as a sum of its discrete and absolutely continuous parts, each weighted by its relative contribution to the overall distribution of $\hat{\tau}^2_{\mathrm{DL}}$:

$$f_{\mathrm{DL}}(t;\tau^2) = F_{\mathrm{M}}(0;\tau^2)I_{\{0\}}(t) + [1-F_{\mathrm{M}}(0;\tau^2)]\left[\frac{f_{\mathrm{M}}(t;\tau^2)}{1-F_{\mathrm{M}}(0;\tau^2)}\right]I_{(0,\infty)}(t) \tag{9}$$

where $F_{\mathrm{M}}(\cdot;\tau^2)$ is the cumulative distribution function associated with density $f_{\mathrm{M}}(\cdot;\tau^2)$.[18] Using this distribution, for each $i=1,2,\ldots,K$, we define the weights $w_i^*(\tau^2)$ as

$$\begin{aligned} w_i^*(\tau^2) &= \mathbb{E}[w_i(\hat{\tau}^2_{\mathrm{DL}})] \\ &= F_{\mathrm{M}}(0;\tau^2)w_i(0) + \int_0^\infty w_i(t)f_{\mathrm{M}}(t;\tau^2)\,\mathrm{d}t. \end{aligned} \tag{10}$$

These weights also depend on τ^2, through the functions $F_{\mathrm{M}}(\cdot;\tau^2)$ and $f_{\mathrm{M}}(\cdot;\tau^2)$. As above, when estimating these functions, we substitute $\hat{\tau}^2_{\mathrm{DL}}$ for τ^2 and proceed with the functions so specified. Estimates of the $w_i^*(\tau^2)$ are then $w_i^*(\hat{\tau}^2_{\mathrm{DL}})$.

This leads naturally to a new estimate for μ

$$\hat{\mu}^*(\hat{\tau}^2_{\mathrm{DL}}) = \frac{\sum w_i^*(\hat{\tau}^2_{\mathrm{DL}})\,Y_i}{\sum w_i^*(\hat{\tau}^2_{\mathrm{DL}})} \tag{11}$$

Table I. Diuretic and pre-eclampsia data

Study	OR	95% CI	$Y_i = \log(\text{OR})$	$\hat{\sigma}_i^2$
Weseley	1·04	(0·48, 2·28)	0·042	0·160
Flowers	0·40	(0·20, 0·78)	− 0·924	0·118
Menzies	0·33	(0·14, 0·74)	− 1·122	0·178
Fallis	0·23	(0·08, 0·67)	− 1·473	0·299
Cuadros	0·25	(0·13, 0·48)	− 1·391	0·114
Landesman	0·74	(0·59, 0·94)	− 0·297	0·015
Krans	0·77	(0·39, 1·52)	− 0·262	0·121
Tervila	2·97	(0·59, 15·07)	1·089	0·686
Campbell	1·14	(0·69, 1·91)	0·135	0·068

and an estimated variance

$$\widehat{\text{var}}[\hat{\mu}^*(\hat{\tau}^2_{\text{DL}})] = \frac{1}{(\sum w_i^*(\hat{\tau}^2_{\text{DL}}))^2} \sum [w_i^*(\hat{\tau}^2_{\text{DL}})]^2(\sigma_i^2 + \hat{\tau}^2_{\text{DL}}). \tag{12}$$

We can use the estimates (11) and (12) with an approximating normal distribution in the usual way to obtain CIs for μ.

This method has the benefit that it takes into account the variability of $\hat{\tau}^2_{\text{DL}}$ in a natural way. If τ^2 is estimated tightly, which Larholt *et al.*[15] conclude should be the case if 20 or more studies are combined, then the weights should be similar to those in the standard random effects model. If, however, we have only limited knowledge of τ^2, reflected in a wide spread for $f_{\text{DL}}(\cdot\,; \tau^2)$, then this is also reflected in the new weighting scheme.

Appendix II provides algorithms for implementing these computations in SAS.[14]

We now apply this to two different meta-analyses in which the revised weighting has rather different effects.

4. APPLICATIONS

4.1. Diuretics and pre-eclampsia

Collins *et al.*[8] presented a review of the randomized trials of prevention of pre-eclampsia with diuretics, and Thompson and Pocock[5] applied the RE model for meta-analysis to these data.

Table I lists the individual odds ratios (ORs), corresponding logit-based CIs, Y_is, and logit variance estimates $\hat{\sigma}_i^2$ for this example, taken from reference 5, p. 1128.

Comparison of the individual studies' ORs in column 2 of Table I suggests considerable disparity among the results. The CIs in column 3, however, indicate more agreement than the point estimates alone suggest. While three of the point estimates are greater than unity, in drastic contrast to the four below 0·50, none is significantly different from unity, as seen by the accompanying CIs. The observed value of the statistic Q in equation (2) is 27·27, with a p-value of 0·001, while the likelihood ratio test results in $\lambda_K(0) = 6{\cdot}39$ with a p-value of 0·011, both indicating heterogeneity.

Table II contains the results of application of the methods developed in Sections 2.1 and 2.2 for inference on τ^2 to this diuretic and pre-eclampsia data. Both moment and maximum likelihood methods of point estimation agree closely and give positive estimates. The AML method yields an upper limit of 0·57 for τ^2, substantially smaller than the other two results; the ALR method gives

Table II. Results for diuretic and pre-eclampsia data: estimation of τ^2

Method	$\hat{\tau}^2$	$(\hat{\tau}^2_L, \hat{\tau}^2_U)$
MM	0·23	(0·04, 2·35)
ALR	0·24	(0·03, 1·13)
AML	0·24	[0, 0·57)

Table III. Results for diuretic and pre-eclampsia data: estimation of $\rho = \exp(\mu)$

Method	$\hat{\rho}(\hat{\tau}^2)$	95% CI
Fixed effects	0·67	(0·56, 0·80)
Random effects (both standard methods)	0·60	(0·40, 0·89)
Random effects (accounting for variability in $\hat{\tau}^2_{\mathrm{DL}}$)	0·62	(0·41, 0·96)

an intermediate upper limit of 1·13 for τ^2 compared with the upper limit of 2·35 from the MM method. Both the MM and ALR CIs significantly indicate a positive τ^2.

That the MM method gives the widest of the three CIs is not unexpected, since there are relatively few studies combined and the asymptotic methods tend to be optimistic with small sample sizes. The disparity between the two maximum likelihood approaches is unexpected, though a reasonable explanation is that the approximating normal and χ^2 distributions differ so greatly as a result of the skewness of the χ^2 distribution and the relatively small number of studies; the non-truncated value for the lower CI limit with the AML method was $-0{\cdot}10$, markedly less than the 0·03 with the ALR method.

Table III gives the standard RE point and 95 per cent CIs for $\rho = \exp(\mu)$ using the usual point estimators for τ^2. We provide the analogous FE estimates for comparison. The functional notation $\hat{\rho}(\hat{\tau}^2)$ in the column heading is to indicate the use of the point estimate for τ^2 in equations (3) and (4) when computing $\hat{\rho}(\hat{\tau}^2) = \exp\{\hat{\mu}(\hat{\tau}^2)\}$, with the CIs determined analogously.

The standard RE point estimates and CIs for ρ agree for all three methods; this follows from the agreement of the point estimates for τ^2 in Table II. The reasonably large values of $\hat{\tau}^2$ have given more weight to the smaller studies with very small ORs (those of Fallis and Menzies) when estimating ρ, resulting in a decrease in the estimates for ρ from 0·67 with the FE model to 0·60 for the RE methods. Note that the Larholt *et al.*[15] approximation to var(Q) in equation (8) is inappropriate since the values of $\hat{\tau}^2$ are actually smaller than two of the $\hat{\sigma}_i^2$. While each of the CIs for ρ is significant at the 95 per cent level, that is, does not contain the null value $\rho = 1{\cdot}0$, the RE CIs are considerably wider than the corresponding FE CI. In general, one would expect this, since, by including a positive value for the parameter τ^2, we are acknowledging greater variation in the observations and consequently greater uncertainty when estimating ρ.

Using the weights $w_i^*(\hat{\tau}^2_{\mathrm{DL}})$ from equation (10) in the formulae (11) and (12) produces the last line of Table III. The point estimate 0·62 is slightly higher than the 0·60 of the other two methods, and the CI is wider, with the upper limit 0·96 noticeably higher than the upper limit 0·89 given by the standard techniques. As our method takes into account variation in the estimate for τ^2, a wider CI is not unexpected. Our proposed approach yields an interval much closer to statistical insignificance (at 95 per cent) than the standard methods suggest. Indeed, such variation might in

Table IV. Relative weights (per cent) for diuretic and pre-eclampsia studies using weights w_i, $w_i(\hat{\tau}^2_{DL})$, and $w_i^*(\hat{\tau}^2_{DL})$

Study	w_i	$w_i(\hat{\tau}^2_{DL})$	$w_i^*(\hat{\tau}^2_{DL})$
Weseley	5·06	10·66	8·42
Flowers	6·86	11·94	10·11
Menzies	4·55	10·19	7·88
Fallis	2·71	7·86	5·59
Cuadros	7·10	12·08	10·31
Landesman	53·96	16·95	30·97
Krans	6·69	11·84	9·96
Tervila	1·18	4·54	3·00
Campbell	11·90	13·94	13·78

some cases change conclusions drawn from application of the standard techniques; if one adopted, as one might well, a one-sided 99 per cent CI to test if the use of diuretics is beneficial, then with the usual RE model one computes an upper confidence bound of 0.96, thus rejecting the hypothesis of no effect at this level; but if one uses our weights w_i^* incorporating the distribution of τ^2, then one gets an upper confidence bound of 1·03, and the hypothesis of no effect is not rejected at this level.

One concern raised about the DerSimonian–Laird method when $\hat{\tau}^2_{DL}$ is large is that, by making the weights close to uniform, it gives too much weight to relatively small studies (with large variances σ_i^2), or equivalently too little weight to the larger studies;[5] that is, their method over-compensates for across-study variability. It is instructive to see how our weights $w_i^*(\hat{\tau}^2_{DL})$ differ from the DerSimonian–Laird weights $w_i(\hat{\tau}^2_{DL})$ and from the FE weights w_i.

Table IV lists for each of these methods the weights of the individual studies as a percentage of the sum of all the studies' weights, and here we see that our proposed weighting procedure does not go as far as the standard DerSimonian–Laird method in this regard. For example, the FE weight w_i given to the largest study's observation (Landesman) is 53·96 per cent of the total available weight; under the DerSimonian–Laird method the weight for this study reduces to 16·95 per cent; and the weight afforded this study under our weighting scheme is 30·97 per cent. Likewise, the smallest study (Tervila, 1·18 per cent) receives 4·54 per cent of weight by the DerSimonian–Laird method, while the same study receives 3·00 per cent by our method.

4.2. Environmental tobacco smoke and lung cancer

In our second example we consider the relationship between environmental tobacco smoke (ETS) and the incidence of lung cancer. There have been several meta-analyses of this relationship performed, see, for example, references 9, 20–24. We consider here the case-control studies of females exposed to spousal ETS in the home and their relative risk of contracting lung cancer. The data we use come from Table 2 of Mengersen *et al.*,[9] which reports the individual ORs and 95 per cent logit CIs from 32 case-control studies of this association. As there are many studies in this example, we only summarize the data here. The largest OR among these studies is 2·55 and the smallest 0·74. Of the 32 studies, 23 reported ORs greater than unity, and 6 of these were significant at the 95 per cent level (logit method). The largest two studies reported estimates 1·26 [95 per cent CI (1·04, 1·54)] and 0·96 [95 per cent CI (0·77, 1·20)]; their individual study variances were 0·010 and 0·013, respectively.

Table V. Results for ETS and lung cancer data: estimation of τ^2

Method	$\hat{\tau}^2$	$(\hat{\tau}^2_L, \hat{\tau}^2_U)$
MM	0·030	[0, 0·14)
ALR	0·032	(0·003, 0·11)
AML	0·032	[0, 0·08)

The observed value of Q from equation (2) is 46·27, with a p-value of 0·04 for Cochran's test of homogeneity, while the likelihood ratio test resulted in $\lambda_K(0) = 5{\cdot}52$ with a p-value of 0·02. This indication of heterogeneity suggests that the FE is likely to be inappropriate in this instance, especially in view of the low power of Cochran's test. Table V shows point estimates and 95 per cent CIs for τ^2. There is reasonable agreement of the point estimates and associated CIs of τ^2 in this example. This is not unexpected since there are many more studies combined here compared with the previous example, so the asymptotic results should fare better. Note that only the ALR CI for τ^2 does not contain the null value $\tau^2 = 0$, again reflecting the marginal heterogeneity as measured by the tests for homogeneity.

All three CIs for ρ from the RE model agree, since the point estimates for τ^2 agree. The estimate of odds ratio ρ using the weights $w_i^*(\hat{\tau}^2_{\mathrm{DL}})$ and equations (11) and (12) is also in close agreement with the standard methods in this case. In contrast to the preceeding example, the CI here is marginally shorter than for the standard methods. This could be due to minor round-off in the numerical computations, or it might be a reflection of the relatively large number of studies here available, which could yield a better approximation for the distributions $f_{\mathrm{M}}(\cdot\,; \tau^2)$ and $f_{\mathrm{DL}}(\cdot\,; \tau^2)$ and thus potentially better estimates $w_i^*(\tau^2)$ for $w_i(\tau^2)$ than with the standard RE approach.

Computation of the relative weights given the individual studies' results by the different methods resulted in the same pattern seen in the preceeding example. The relatively large studies are down-weighted by the DerSimonian–Laird RE methods much more than by our proposed weighting scheme; and the relatively small studies do not receive as much weight by our proposed weighting scheme as with the DerSimonian–Laird method. For the largest study noted above, the relative weight given by w_i was 18·13 per cent; this reduced to 8·73 per cent with the DerSimonian–Laird method and to 11·43 per cent with our weighting scheme. For the second largest study, these values were 13.95 per cent for the FE weight w_i, and 8·13 per cent and 9·85 per cent for the DerSimonian–Laird and our proposed methods, respectively.

5. COMMENTS

As the two examples show, incorporating the variability in estimates of τ^2 into the overall mean CI can but does not always produce a substantial difference in the inference being carried out. The interval estimate based on our weights for the odds ratio ρ in the ETS/lung cancer example, reported in Table VI, does not differ greatly from the standard methods, although for the diuretic/pre-eclampsia example in Table III the difference is much more noticeable. The proposed weighting scheme ought to prove useful when the number of studies K is fairly small (say less than 20 based on the results of Larholt *et al.*);[15] for larger collections of studies the more major contribution may well be tests of $\tau^2 > 0$, which if performed using the moment-based CI also incorporates the variability in $\hat{\tau}^2_{\mathrm{DL}}$. Note also that since $w_i(\tau^2) \approx 1/\tau^2$ when $\tau^2 \gg \sigma_i^2$, and

Table VI. Results for ETS and lung cancer data: estimation of $\rho = \exp(\mu)$

Method	$\hat{\rho}(\hat{\tau}^2)$	95% CI
Fixed effects	1·15	(1·06, 1·26)
Random effects (both standard methods)	1·20	(1·70, 1·35)
Random effects (accounting for variability in $\hat{\tau}^2_{\mathrm{DL}}$)	1·19	(1·06, 1·34)

$w_i(\tau^2) \approx 1/\sigma_i^2$ when $\sigma_i^2 \gg \tau^2$, we expect the proposed method to provide real improvement when $\hat{\tau}^2_{\mathrm{DL}}$ is comparable to at least some of the σ_i^2.

We observe in the examples in Section 4 that our method may provide a balance between the two standard methods (the FE and the DerSimonian–Laird approaches). One can show that in general $w_i(0) \geqslant w_i^*(\hat{\tau}^2_{\mathrm{DL}}) \geqslant w_i(\hat{\tau}^2_{\mathrm{DL}})$ (at least approximately), using the fact that $w_i(t)$ is monotone decreasing in t together with the convexity of $w_i(t)$ and Jensen's inequality. In the examples our relative weights also fell between the FE and DerSimonian–Laird relative weights; this is not to be expected in general.

In this paper we have approached the RE model in meta-analysis wholly from a frequentist perspective. The inability of the current frequentist methodology to handle variability in estimates of τ^2 when testing $\tau^2 > 0$ and when estimating the overall mean μ were primary motivations for the present work.

We note that Bayesian methods in meta-analysis have also been developed by several authors using the same hierarchic structure employed here in equations (1). A consistent development and comparison between frequentist, empirical Bayesian, and full Bayesian methods, indicating the potential impact the choice of method might have, is given in reference 9; several references to the Bayesian approach are given there. Results in the present paper may be naturally incorporated into standard hierarchic Bayesian meta-analysis. In particular, our approximating distribution for $\hat{\tau}^2_{\mathrm{DL}}$ has immediate application in empirical Bayes methodology.

APPENDIX I: PROOF OF THEOREM 1

We shall use

$$\hat{\mu} = \frac{\sum w_i Y_i}{\sum w_i} \sim \mathrm{N}\left(\mu, \left[\frac{1}{\sum w_i} + \frac{\sum w_i^2}{(\sum w_i)^2}\tau^2\right]\right)$$

and the form of $\mathbb{E}(Q)$ in equation (5).

We compute $\mathbb{E}(Q^2)$. Write Q as

$$Q = \sum w_i(Y_i - \mu)^2 - (\sum w_i)(\hat{\mu} - \mu)^2 \tag{13}$$

noting that all summations are from 1 to K. Squaring the expression in equation (13) and taking expected values

$$\begin{aligned}\mathbb{E}(Q^2) = &\sum_i \sum_j w_i w_j \mathbb{E}\{(Y_i - \mu)^2 (Y_j - \mu)^2\} \\ &- 2(\sum w_j) \sum_i w_i \mathbb{E}\{(Y_i - \mu)^2 (\hat{\mu} - \mu)^2\} \\ &+ (\sum w_i)^2 \mathbb{E}\{(\hat{\mu} - \mu)^4\}. \end{aligned} \tag{14}$$

It is convenient to compute each of these components in turn:

$$\sum_i \sum_j w_i w_j \mathbb{E}\{(Y_i-\mu)^2(Y_j-\mu)^2\} = \sum_i w_i^2 3(w_i^{-1}+\tau^2)^2 + \sum\sum_{i\neq j} w_i w_j (w_i^{-1}+\tau^2)(w_j^{-1}+\tau^2)$$

$$= 2\sum_i w_i^2 (w_i^{-1}+\tau^2)^2 + \left\{\sum_i w_i (w_i^{-1}+\tau^2)\right\}^2;$$

$$-2(\textstyle\sum w_j)\sum_i w_i \mathbb{E}\{(Y_i-\mu)^2(\hat{\mu}-\mu)^2\} = -2(\textstyle\sum w_j)\sum_i w_i \mathbb{E}\left\{(Y_i-\mu)^2\left(\frac{\sum w_j Y_j}{\sum w_j}-\mu\right)^2\right\}$$

$$= -2(\textstyle\sum w_j)\sum_i w_i \mathbb{E}\left\{(Y_i-\mu)^2\left(\frac{\sum w_j (Y_j-\mu)}{\sum w_j}\right)^2\right\}$$

$$= \left(\frac{-2}{\sum w_j}\right)\sum_i w_i \sum_j \sum_k w_j w_k \mathbb{E}\{(Y_i-\mu)^2(Y_j-\mu)(Y_k-\mu)\}$$

$$= \left(\frac{-2}{\sum w_j}\right)\sum_i w_i \left\{3w_i^2(w_i^{-1}+\tau^2)^2 + (w_i^{-1}+\tau^2)\sum_{j\neq i} w_j^2 (w_j^{-1}+\tau^2)\right\}$$

$$= \left(\frac{-2}{\sum w_j}\right)\sum_i w_i \left\{2w_i^2(w_i^{-1}+\tau^2)^2 + (w_i^{-1}+\tau^2)\sum_{j} w_j^2 (w_j^{-1}+\tau^2)\right\}$$

$$= \frac{-2}{\sum w_j}\left\{2\sum_i w_i^3 (w_i^{-1}+\tau^2)^2 + (\textstyle\sum w_i (w_i^{-1}+\tau^2))(\sum w_j^2 (w_j^{-1}+\tau^2))\right\};$$

$$(\textstyle\sum w_i)^2 \mathbb{E}\{(\hat{\mu}-\mu)^4\} = 3(\sum w_i)^2\left\{\frac{1}{\sum w_i} + \tau^2\frac{\sum w_i^2}{(\sum w_i)^2}\right\}^2$$

$$= 3\left(1+\tau^2\frac{\sum w_i^2}{\sum w_i}\right)^2.$$

Substituting these expressions into equation (14) and taking S_1, S_2, and S_3 as in the statement of the theorem, the expression for $\mathbb{E}(Q^2)$ simplifies after some tedious algebra to

$$\mathbb{E}(Q^2) = (K^2-1) + 2(K+1)\left(S_1 - \frac{S_2}{S_1}\right)\tau^2 + \left(S_1^2 - 4\frac{S_3}{S_1} + 3\frac{S_2^2}{S_1^2}\right)\tau^4. \tag{15}$$

Combining equation (15) with equation (5) finally gives, as required,

$$\mathrm{var}(Q) = \mathbb{E}(Q^2) - \{\mathbb{E}(Q)\}^2$$

$$= \mathbb{E}(Q^2) - \left\{(K-1)^2 + 2(K-1)\left(S_1 - \frac{S_2}{S_1}\right)\tau^2 + \left(S_1 - \frac{S_2}{S_1}\right)^2\tau^4\right\}$$

$$= (K^2 - 1) - (K - 1)^2 + \left\{2(K + 1)S_1 - 2(K + 1)\frac{S_2}{S_1} - 2(K - 1)\left(S_1 - \frac{S_2}{S_1}\right)\right\}\tau^2$$
$$+ \left\{S_1^2 - 4\frac{S_3}{S_1} + 3\frac{S_2^2}{S_1^2} - \left(S_1^2 - 2S_2 + \frac{S_2^2}{S_1^2}\right)\right\}\tau^4$$
$$= 2(K - 1) + 4\left(S_1 - \frac{S_2}{S_1}\right)\tau^2 + 2\left(S_2 - 2\frac{S_3}{S_1} + \frac{S_2^2}{S_1^2}\right)\tau^4.$$

APPENDIX II: COMPUTATIONS

At the suggestion of a referee, we provide a guide to performing these computations in the statistical computing package SAS, a registered trademark of the SAS Institute, Inc., Cary, N. C. Similar routines have also been coded by the first author in C and are available from him upon request: the presentation of the procedures in this section is not intended as an endorsement of the SAS Institute or its products.

Users of these routines are cautioned that, while due care has been taken and they are believed accurate, they have not been rigorously tested and their use and results are solely the responsibilities of the user.

Moment CI Limits for τ^2

Computation of the moment interval for τ^2 given in equation (9) requires the use of the gamma probability distribution function, also called the incomplete gamma function. Implementation of this algorithm in SAS might take the following form, which uses the SAS function PROBGAM.[14]

Presume the data input to SAS are variables labelled Y and Sigma2, representing Y_i and σ_i^2 from equations (1). Then, using SAS data steps, compute the values of K, w_i, S_1, S_2, S_3, c, $\hat{\mu}$, Q, $\hat{\tau}^2_{\mathrm{M}}$, and $\hat{\tau}^2_{\mathrm{DL}}$ of Section 2.1. These might be stored in a SAS dataset called sums. The following SAS data step yields the SAS data set MMCI which contains the values t2m $= \hat{\tau}^2_{\mathrm{M}}$, t2dl $= \hat{\tau}^2_{\mathrm{DL}}$, LowerT2 $= \hat{\tau}^2_l$, and UpperT2 $= \hat{\tau}^2_u$. The present algorithm produces an equal-tails, 95 per cent CI.

```
data MMCI(keep = t2m t2dl LowerT2 UpperT2);
  merge sums;

  /* Constants which may need adjusting in any given application  */
  STEP = 0·0001;  /* step size for looping over values of t  */
  MAXERROR = 0·00001;  /* max. error of equations L − 0·025 = 0  */
                                        /*  and U − 0·025 = 0  */
  MAXT2 = 20·0;  /* a value to cut off the number of interactions  */

/*  Loop to compute lower CI limit  */
  Start = − ((K − 1)/C) + STEP;
  do t = Start to MAXT2 by STEP until(ERROR < MAXERROR);
    lambda = ((K − 1) + c*t)/(2*(K − 1) + 4*c*t + 2*(S2 − 2*(S3/S1)
      + ((S2*S2)/(S1*S1)))*t*t);
      r = (K − 1 + c*t)*lambda;
      limit = lambda*(c*t2m + K − 1);
      ERROR = abs((1-probgam(limit, r)) − 0·025);
      /* 0·025 for equal-tails, 95% interval  */
  end;
```



```
LowerT2 = 0<>t;  /*  set lower limit to 0 if it evaluates negative */

/*  Loop to compute upper CI limit  */
Start = LowerT2 + STEP;
ERROR = 1·0;
do t = Start to MAXT2 by STEP until (ERROR < MAXERROR);
  lambda = ((K - 1) + c*t)/(2*(K - 1) + 4*c*t + 2*(S2 - 2*(S3/S1)
    + ((S2*S2)/(S1*S1)))*t*t);
  r = (K - 1 + c*t)*lambda;
  limit = lambda*(c*t2m + K - 1);
  ERROR = abs(probgam(limit, r) - 0·025);
    /* 0·025 for equal-tails, 95% interval  */
end;
UpperT2 = t;
run;
```

The values t2m, t2dl, LowerT2, and Upper T2 are retained in the SAS data set MMCI. Values for STEP, MAXERROR, and MAXT2 should be tailored to the application. For example, the default values listed here were used in the application in Section 4.1, while for the example in Section 4.2 the values were STEP = 0·0001, MAXERROR = 0·001, with MAXT2 the same.

Computation of $w_i^*(\hat{\tau}^2_{DL})$

Computation of $w_i^*(\hat{\tau}^2_{DL})$ in SAS[14] requires the use of PROC IML,[25] which contains the function QUAD[26] to perform numeric integration. Also used is the SAS function PROBGAM.[14] The following code fragment creates two SAS data sets: WSTAR containing the variable wstar = $w_i^*(\hat{\tau}^2_{DL})$; and MUWSTAR containing muhatstar = $\hat{\mu}^*(\hat{\tau}^2_{DL})$ and varmuhatstar = $\widehat{\text{var}}(\hat{\mu}^*(\hat{\tau}^2_{DL}))$.

The input SAS data set obs contains the observations Y and Sigma2. The input SAS dataset sums contains the quantities K, c, $\hat{\tau}^2_M$, S_1, S_2 and S_3.

```
proc iml;
  use obs var{y sigma2};
  read all var {y sigma2};
  use sums var{K C t2m S1 S2 S3};
  read all var{K C t2m S1 S2 S3};

  lambda = ((K - 1) + C*t2m)/(2*(K - 1) + 4*C*t2m + 2*(S2 - 2*(S3/S1)
   + ((S2*S2)/(S1*S1)))*t2m*t2m);
  r = (K - 1 + C*t2m)*lambda;
  w = 1/sigma2;

  /* function to integrate in Equation (10) */
  start Intgrnd(t) global(s, lambda, r, C, K);
    ans = C*(1/(s + t))*
          (lambda**r/gamma(r))*(C*t + K - 1)**(r - 1)*
          exp( - lambda*(C*t + K - 1));
    return(ans);
  finish Intgrnd;

  /* Initalize wStar and scaledwStar with something the right size */
  wStar = sigma2;
  scaledwStar = sigma2;
```



```
/* Compute integral part of Equation (10) for each wStar */
  do i = 1 to K;
    s = sigma2[i, 1];
    a = {0   .P};                    /*  integrate from 0 to infinity  */
    call quad(value, "Intgrnd", a);  /* return evaluation is ''value''  */
    wStar[i, 1] = value;
  end;

  /*  Initialize variables for iterative summation  */
  totwStar = 0;
  VarStar = 0;
  MuHatStar = 0;

  /*  Compute complete wStar values from Equation (10)  */
  /*  and also MuHatStar and VarStar  */
  do i = 1 to K;
    wStar[i, 1] = w[i, 1]*probgam(lambda*(K - 1), r) + wStar[i, 1];
    totwStar = totwStar + wStar[i, 1];
    MuHatStar = MuHatStar + (wStar[i, 1]*y[i, 1]);
    VarStar = VarStar + wStar[i, 1]*wStar[i, 1]*(sigma2[i, 1] + t2m);
  end;

  MuHatStar = MuHatStar/totwStar;
  VarStar = VarStar/(totwStar*totwStar);

  /*  Create SAS dataset containing variable wStar  */
  create WSTAR var{wStar};
    append;
  close WSTAR;

  /*  Create SAS dataset containing variables MuHatStar and VarStar  */
  create MUWSTAR var{MuHatStar VarStar};
    append;
  close MUWSTAR;
```

ACKNOWLEDGEMENTS

We wish to thank Drs. Hariharan K. Iyer and Duane C. Boes of Colorado State University for their comments and suggestions, and also Michael Fiola of Hoechst Marion Roussel for his assistance with SAS. We are also grateful for the helpful comments of Dr. Rebecca DerSimonian, the editor, and an anonymous referee; these strengthened the paper considerably.

Part of this work is included in the Colorado State University Ph.D. dissertation of the first author, with the paper being completed while the same author held a NRC/NIST Postdoctoral Associateship. This paper is a contribution of the National Institute of Standards and Technology and is not subject to copyright in the United States.

The paper was completed with partial support from two tobacco companies; the methods, analysis and conclusions here are, however, entirely those of the authors and should not be otherwise ascribed.

REFERENCES

1. Olkin, I. 'Meta-analysis: methods for combining independent studies', *Statistical Science*, **7**, 226 (1992).
2. Cooper, H. and Hedges, L. V. (eds). *The Handbook of Research Synthesis*, Russell Sage Foundation, New York, 1994.
3. NRC Committee on Applied and Theoretical Statistics. *Combining Information*: *Statistical Issues and Opportunities for Research*, National Academy Press, Washington, 1992.
4. Hedges, L. V. and Olkin, I. *Statistical Methods for Meta-analysis*. Academic Press, 1985.
5. Thompson, S. G. and Pocock, S. J. 'Can meta-analyses be trusted?', *Lancet*, **338**, 1127–1130 (1991).
6. Greenland, S. 'A critical look at some popular meta-analytic methods', Unpublished, 1993.
7. DerSimonian, R. and Laird, N. M. 'Meta-analysis in clinical trials', *Controlled Clinical Trials*, **7**, 177–188 (1986).
8. Collins, R., Yusuf, S. and Peto, R. 'Overview of randomized trials of diuretics in pregnancy', *British Medical Journal*, **290**, 17–23 (1985).
9. Mengersen, K. L., Tweedie, R. L. and Biggerstaff, B. J. 'The impact of method choice in meta-analysis', *Australian Journal of Statistics*, **37**, 19–44 (1995).
10. Cochran, W. G. 'The combination of estimates from different experiments', *Biometrics*, **10**, 101–129 (1954).
11. Raudenbush, S. W. 'Random effects models', *in* Cooper, H. and Hedges, L. V. (eds), *The Handbook of Research Synthesis*, Russell Sage Foundation, 1994, pp. 301–321.
12. Waternaux, C. and DuMouchel, W. 'Combining information across sites with hierarchical Bayesian linear models', In *Proceedings of the Section on Bayesian Statistics*. ASA, 1993. Joint Statistical Meetings, San Francisco.
13. DuMouchel, W. *Bayesian Meta-analysis*, Berry, D. (ed), Marcel Dekker, New York, 1990, pp. 509–529.
14. The SAS Institute Inc. *SAS Language: Reference, Version 6*, *First Edition*. The SAS Institute, Inc., Cary, NC, 1990.
15. Larholt, K. M., Tsiatis, A. A. and Gelber, R. D. 'Variability of coverage probabilities when applying a random effects methodology for meta-analysis', Unpublished, 1994.
16. Satterthwaite, F. E. 'An approximate distribution of estimates of variance components', *Biometrics Bulletin*, 110–114 (1946).
17. Rao, P. S. R. S., Kaplan, J. and Cochran, W. G. 'Estimators for the one-way random effects model with unequal error variances, *Journal of the American Statistical Association*, **76**, (373), 89–97 (1981).
18. Mood, F. A., Graybill and Boes, D. C. *Introduction to the Theory of Statistics*, 3rd edn, McGraw-Hill, New York, 1963.
19. Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. *Numerical Recipes in C*, 2nd edn, Cambridge University Press, Cambridge, UK, 1992.
20. EPA Draft Review. *Health Effects of Passive Smoking: Assessment of Lung Cancer in Adults and Respiratory Disorders in Children*. National Academy Press, United States EPA, Washington, 1990.
21. EPA Review, *Health Effects of Passive Smoking: Assessment of Lung Cancer in Adults and Respiratory Disorders in Children*. National Academy Press, United States EPA, Washington, 1992.
22. Biggerstaff, B. J., Mengersen, K. L. and Tweedie, R. L. 'Passive smoking in the workplace: Classical and Bayesian meta-analyses', *International Archives of Occupational and Environmental Health*, **66**, 269–277 (1994).
23. Lee, P. N. *An Assessment of the Epidemiological Evidence Relating Lung Cancer Risk in Never Smokers to Environmental Tobacco Smoke Exposure*, Kasuga, H. (ed), Springer-Verlag, New York, 1993, pp. 28–70.
24. Lee, P. N. 'Passive smoking', *Fd. Chem. Toxic.*, **20**, 223–229 (1981).
25. The SAS Institute Inc. *SAS/IML Software*: *Usage and Reference*, *Version 6*, *First Edition*, The SAS Institute, Inc., Cary, NC, 1989.
26. The SAS Institute Inc. *SAS Technical Report P-230*, *SAS/IML Software*: *Changes and Enhancements*, *Release 6.07*, The SAS Institute, Inc., Cary, NC, 1991.

