Tatra Mt. Math. Publ. **26** (2003), 155–168

TATRA MOUNTAINS
Mathematical Publications

TWO PROCEDURES OF META-ANALYSIS IN CLINICAL TRIALS AND INTERLABORATORY STUDIES

Andrew L. Rukhin

ABSTRACT. The random effects model designed to combine information from several clinical trials is related to an analysis of data from interlaboratory studies. The primary goal of this work is to establish the relationship between the estimator proposed by DerSimonian and Laird [*Meta-analysis in clinical trials* Controlled Clinical Trials **7** (1986), 177–188] and the Mandel–Paule method [*Consensus values and weighting factors* J. Res. Nat. Bureau Standards **87** (1982), 377–385] used in collaborative studies. Both procedures estimate the overall treatment effect via a moment-type estimator and give approximate formulas for their standard errors. Theoretical properties of these estimates are compared in the assymptotic setting. To this end, a class of weighted means statistics is investigated for situations where the number of sources is large. It is shown that both the Mandel–Paule rule and the DerSimonian–Laird estimator of the between-trials variance are inconsistent in this setting. The results of numerical comparison of mean squared errors of these estimators for a special distribution of within-trials variances are also reported.

1. Statement of the problem

Statistical analysis of many multicenter clinical trials and that of interlaboratory studies exhibit many similarities. The latter are used when certifying, for example, standard reference materials, the former when a comparison between two treatments is to be established. Both of these areas commonly use a random effects model as a statistical tool. Actually, it becomes the model of choice in the more general situation where one has to combine information from several comparative studies.

2000 Mathematics Subject Classification: Primary 62F12; Secondary 62E20, 62J05, 62N99.
Keywords: clinical trials, consistency, DerSimonian–Laird estimator heteroscedasticity, interlaboratory study, Mandel–Paule algorithm, maximum likelihood estimator, meta-analysis, random effects model.

We assume that the data in the ith trial/laboratory has an additive error structure consisting of a random laboratory effect b_i, and of the measurement errors e_{ij}. More precisely, it is assumed that the data x_{ij} has the form

$$x_{ij} = \mu + b_i + \varepsilon_{ij} . \tag{1}$$

Here $i = 1, \ldots, p$ indexes the laboratories, and $j = 1, \ldots, n_i$ represents the sample size (the number of measurements) in laboratory i; μ is the true mean (reference value). The random variables b_i and ε_{ij} are all independent and normal with zero means and variances σ_B^2 and τ_i^2. The random variables b_i represent the between-study effect which is commonly observed in collaborative studies. It is possible that in (1) $b_i \equiv 0$ and $\sigma_B^2 = 0$.

Clearly, (1) leads to the following model for the sample means $x_i = \bar{x}_i = \sum x_{ij}/n_i$,

$$x_i = \mu + b_i + e_i . \tag{2}$$

Here mutually independent random variables are normally distributed, $b_i \sim N(0, \sigma_B^2)$ and $e_i \sim N(0, \sigma_i^2)$. In this situation (unbiased) estimates, $s_i^2 = \sum_j (x_{ij} - x_i)^2 / \left[n_i(n_i - 1)\right]$, of the within-trials variances $\sigma_i^2 = \tau_i^2/n_i$ are available.

A fundamental problem in meta-analysis is to estimate the overall treatment effect μ, and to provide a standard error for this estimate. In two-arm multiple clinical trials, if the data in each trial can be assumed to be normal, x_i represents the difference between the mean responses in this trial, and s_i^2 is the within-trial sample variance. If the data is binary, then x_i is the logarithm of odds ratio $\log\left(p_{1i}(1 - p_{2i})/[p_{2i}(1 - p_{2i})]\right)$, and (2) holds approximately with $\sigma_i^2 = \left[p_{1i}(1 - p_{1i})\right]^{-1} + \left[p_{2i}(1 - p_{2i})\right]^{-1}$.

DerSimonian and Laird (1986) suggested an estimator of μ, which turned out to be very popular in the meta-analysis of multicenter clinical trials. In fact, the number of references to the paper by DerSimonian and Laird (1986) exceeds 1250. Partly, this popularity is due to the fact that this is a simple non-iterative procedure, which admits an approximate formula for the variance of the resulting estimator.

The model (2) also appears in interlaboratory studies where μ represents the common effect, or consensus mean. In this case, x_i is the sample mean of the data obtained by ith laboratory, and s_i^2 is the rescaled sample variance which estimates the variance of x_i. See Mandel (1991) or Crowder (1992) for a detailed discussion. Cochran (1937, 1954) and Rao (1981) discuss the classical maximum likelihood estimate of μ. Because of the rather complicated form of the likelihood equations, simpler procedures are desirable in practice. A method for estimating the common mean μ was suggested by Mandel and Paule (1970) and in a more general form by Paule and Mandel (1982). It is now widely used in applications, particularly in analytical chemistry. Experience has shown that this method, the *Mandel–Paule* algorithm, often provides reasonable estimates. It is recommended (Schiller and Eberhardt, 1992) for use in the analysis of reference materials in international metrological settings. This estimator can been seen to be a version of the approximate maximum likelihood estimator (Rukhin and Vangel, 1998).

Both of these methods are discussed in Section 2. Some theoretical properties of these methods are derived in Sections 3 and 4 for a large number of centers (laboratories). One of the goals of this work is to caution about the use of estimators of the between-study variance, and of the mean squared errors of these two procedures.

The author is grateful to the referee for helpful comments. Thanks are also due to Stefan Leigh and Mark Vangel for many interesting discussions.

2. Weighted means statistics

When the within-trials and between-trials variances σ_i^2 and σ_B^2 are known, the best (in terms of the mean squared error) unbiased estimator of the treatment effect μ in the model (2) is a weighted means statistic

$$\tilde{x} = \frac{\sum_{i=1}^{p} w_i x_i}{\sum_{i=1}^{p} w_i} \tag{3}$$

with $w_i = w_i^0 = \dfrac{1}{\sigma_B^2 + \sigma_i^2}$. Even without the normality assumption for these optimal weights,

$$E \sum_i w_i^0 (x_i - \tilde{x})^2 = p - 1\,,$$

and

$$E(\tilde{x} - \mu)^2 = \frac{1}{\sum_i w_i^0} = \frac{1}{\sum_i (\sigma_B^2 + \sigma_i^2)^{-1}}\,. \tag{4}$$

If $\mathrm{Var}(x_i) = \sigma_i^2 + \sigma_B^2$, but the weights w_i are arbitrary,

$$\begin{aligned} E \sum_i w_i (x_i - \tilde{x})^2 &= \sum_1^p (\sigma_B^2 + \sigma_i^2) w_i - \frac{\sum_1^p (\sigma_B^2 + \sigma_i^2) w_i^2}{\sum_1^p w_i} \\ &= \sigma_B^2 \left[\sum_1^p w_i - \frac{\sum_1^p w_i^2}{\sum_1^p w_i} \right] + \sum_1^p \sigma_i^2 w_i - \frac{\sum_1^p \sigma_i^2 w_i^2}{\sum_1^p w_i}\,. \end{aligned} \tag{5}$$

In particular, when $w_i = 1/\sigma_i^2$

$$E\sum_i \frac{(x_i - \tilde{x})^2}{\sigma_i^2} = p - 1 + \sigma_B^2 \left[\sum_1^p \frac{1}{\sigma_i^2} - \frac{\sum_1^p \frac{1}{\sigma_i^4}}{\sum_1^p \frac{1}{\sigma_i^2}} \right]. \tag{6}$$

It makes sense to estimate the within-trials variances σ_i^2 by the available estimates s_i^2, but the problem of estimating the between-study component of variance σ_B^2 remains.

By employing the idea behind the method of moments, DerSimonian and Laird (1986) made use of the identity (6) as an estimating equation for μ and σ_B^2, provided that σ_i^2 are estimated by s_i^2, in the following way. For weights of the form

$$w_i = \frac{1}{y + s_i^2}, \tag{7}$$

determine a non-negative $y = y_{\mathrm{DL}}$ from the formula

$$y_{\mathrm{DL}} = \max \left[0, \frac{\sum_i s_i^{-2}(x_i - \tilde{x}^0)^2 - p + 1}{\sum_{i=1}^p s_i^{-2} - \sum_{i=1}^p s_i^{-4} \left[\sum_{i=1}^p s_i^{-2} \right]^{-1}} \right]. \tag{8}$$

Here

$$\tilde{x}^0 = \frac{\sum_{i=1}^p x_i s_i^{-2}}{\sum_{i=1}^p s_i^{-2}}$$

is one of traditional estimators of the common mean originally suggested by Graybill and Deal (1959). In other terms, the statistic $\tilde{x}^0$ and the weights $w_i = s_i^{-2}$, corresponding to $\sigma_B^2 = 0$, are used to evaluate the sum $\sum_i s_i^{-2}(x_i - \tilde{x}^0)^2$, which is employed to estimate the true σ_B^2 via (6).

DerSimonian and Laird, motivated by (4), also gave an approximate formula for the estimate of the variance of the resulting estimator,

$$\widehat{\mathrm{Var}}(\tilde{x}) = \frac{1}{\sum_i (y_{\mathrm{DL}} + s_i^2)^{-1}}.$$

In another application of (5), take $w_i = 1/p$, so that $\tilde{x} = \bar{x}$. The identity

$$\frac{E\sum_{i=1}^p (x_i - \bar{x})^2}{p-1} = \sigma_B^2 + \frac{\sum_{i=1}^p \sigma_i^2}{p},$$

leads to a classical ANOVA-type estimate of σ_B^2,

$$y_{\mathrm{AN}} = \max\left[0, \frac{\sum_{i=1}^{p}(x_i - \bar{x})^2}{p-1} - \frac{\sum_{i=1}^{p} s_i^2}{p}\right].$$

The Mandel–Paule algorithm uses weights of the form (7) as well, but now $y = y_{\mathrm{MP}}$, which is designed to approximate σ_B^2, is found from the moment-type estimating equation

$$\sum_{i=1}^{p} \frac{(x_i - \tilde{x})^2}{y_{\mathrm{MP}} + s_i^2} = p - 1\,. \tag{9}$$

The *modified Mandel–Paule* procedure with $y = y_{\mathrm{MMP}}$ is defined as above, but $p-1$ in the right-hand side of (9) is replaced by p, i.e.,

$$\sum_{i=1}^{p} \frac{(x_i - \tilde{x})^2}{y_{\mathrm{MMP}} + s_i^2} = p\,. \tag{10}$$

The DerSimonian–Laird procedure and both Mandel–Paule rules provide estimates $\tilde{x}$ of the consensus mean μ along with the value y which is commonly used to estimate σ_B^2. The first estimates are quite satisfactory, but the latter lack some desirable properties.

Notice that when $p = 2$, the DerSimonian–Laird estimator coincides with the Mandel–Paule rule, as, in this case,

$$y_{\mathrm{MP}} = y_{\mathrm{DL}} = \frac{1}{2}\max\left[0, (x_1 - x_2)^2 - s_1^2 - s_2^2\right].$$

In the general case, both of these rules set $y = 0$ when

$$\sum_{i=1}^{p} \frac{(x_i - \tilde{x})^2}{s_i^2} \le p - 1\,.$$

Actually, the DerSimonian–Laird procedure is a one-step iteration designed to solve (9) written in the form suggested by (5), with the initial value $y = 0$,

$$\sum_{i=1}^{p} \frac{(x_i - \tilde{x})^2}{y + s_i^2} - \sum_{i=1}^{p} \frac{s_i^2}{y + s_i^2}\left[1 - \frac{1}{y + s_i^2}\left(\sum_{i=1}^{p}\frac{1}{y + s_i^2}\right)^{-1}\right]$$
$$= y\left[\sum_{i=1}^{p}\frac{1}{y + s_i^2} - \sum_{i=1}^{p}\frac{1}{(y + s_i^2)^2}\left(\sum_{i=1}^{p}\frac{1}{y + s_i^2}\right)^{-1}\right].$$

It was shown by Rukhin and Vangel (1998) that the modified Mandel–Paule rule is characterized by the following fact: the maximum likelihood estimator $\hat{\sigma}_B^2$ of σ_B^2 coincides with y_{MMP}, if in the reparametrized version of the likelihood equation the weights w_i admit representation (7). Section 3 of Rukhin, Biggerstaff and Vangel (2000) gives a bound on the difference between the usual Mandel–Paule solution and the restricted maximum likelihood estimator.

As a consequence, the corresponding weighted means statistic (3) must be close to the maximum likelihood estimator, so that both versions of the modified Mandel–Paule estimator approximate their maximum likelihood counterparts. Thus, the modified Mandel–Paule estimator can be interpreted as a procedure which uses weights of the form $1/(y+s_i^2)$, instead of solutions of the likelihood equation that are difficult to find, and still maintains the same estimate of σ_B^2 as the maximum likelihood. A similar interpretation holds for the original Mandel–Paule rule and the restricted likelihood function. For this reason both Mandel–Paule rules are quite natural. It is also suggestive to use the weights (7) with $y = y_{\mathrm{MP}}$ determined from the Mandel–Paule equation (9) as a first approximation when solving the likelihood equations.

3. Asymptotic behavior of weighted means

In this Section we look at the asymptotic behavior of the class of statistics that includes the modified maximum likelihood procedure and the Mandel–Paule rule.

This class is composed of general weighted means statistics $\tilde{x}$ of the form (3) with w_i given by (7). The value of y is determined from an estimating equation such as (8) or (9). As it happens, under the assumptions detailed below this value converges with probability one to a constant obtained from the limiting form of the estimating equations. For this reason we study the behavior of statistics (3) for a fixed positive y.

To perform an asymptotic comparison of these statistics we assume that $p \to \infty$. In many collaborative studies it is believed that despite unbalancedness, the within-trials variances, σ_i^2, can be regarded as realizations of a random variable, and we will assume that σ_i^2 are i.i.d. replicas of a random variable with some fixed but otherwise arbitrary distribution function $G(\cdot)$. Although in practice the elicitation of the form of G from practitioners is very difficult, this setting is useful since asymptotic variance estimation for the statistics (3) becomes possible. The asymptotic variance so obtained agrees with estimates derived by other methods as well as with simulation results. An alternative condition would be that the σ_i^2 can be non-identically distributed, but such that an analogue of the Lindeberg condition for the independent random variables $x_i/(y+s_i^2)$ is satisfied.

Let the observable i.i.d. random variables x_i, s_i^2, $i = 1, 2, \ldots, p$, be realizations of the random vector (X, S^2) such that X and S^2 are conditionally (for given σ) independent with the conditional distribution of X being $N\left(\mu, \sigma_B^2 + \sigma^2\right)$ for some unknown σ_B^2. For simplicity, we take both μ and σ_B^2 to be fixed. A similar model has been used by Cox (1975) in the situation when $\sigma_B^2 = 0$; ours is the simplified version of the setting in Rukhin and Vangel (1998).

If it is desired to have a fully Bayesian model, σ_B^2 and μ can be assumed to be random variables, in which case the conditional distributions of X and S^2 are those for given σ_B^2, σ^2 and μ. The conditional distribution of S^2 is supposed to be of the form $\sigma^2 W$ with a random variable W, $EW = 1$, having known distribution, say, with density h. In the typical (Gaussian) case employed in the simulations in Section 4, W has the distribution of χ_ν^2/ν with ν being the typical degrees of freedom (or a mixture of such distributions); random variables W, σ^2 are independent.

Thus, $E(X|\sigma) = \mu$, and $E\big([X-\mu]^2|\sigma\big) = \sigma_B^2 + \sigma^2$. The joint density of the random vector (X, S) is then

$$\frac{1}{\sqrt{2\pi}} \int \exp\left\{-\frac{(x-\mu)^2}{2(\sigma^2+\sigma_B^2)}\right\} \frac{h(s/\sigma)}{\sigma\sqrt{\sigma_B^2+\sigma^2}}\, dG(\sigma).$$

The law of large numbers shows that for a fixed y,

$$\frac{1}{p}\sum_1^p \frac{1}{y+s_i^2} \to E\frac{1}{y+S^2}$$

and

$$\tilde{x} \to \frac{E\frac{X}{y+S^2}}{E\frac{1}{y+S^2}} = \frac{E\left[E(X|\sigma)E\left(\frac{1}{y+S^2}|\sigma\right)\right]}{E\left[\left(E\frac{1}{y+S^2}|\sigma\right)\right]} = \mu\,.$$

Thus, under our assumptions, $\tilde{x}$ is a consistent estimator of μ, and, according to the Central Limit Theorem, $p^{-1/2}\sum_1^p w_i(x_i-\mu) = p^{-1/2}(\tilde{x}-\mu)\sum_1^p w_i$ has an approximately normal distribution with zero mean and variance $E(X-\mu)^2(y+S^2)^{-2}$. Therefore, $p^{1/2}(\tilde{x}-\mu)$ is asymptotically normally distributed with zero mean and the variance

$$S(y) = \frac{E\frac{(X-\mu)^2}{(y+S^2)^2}}{\left(E\frac{1}{y+S^2}\right)^2} = \frac{E\frac{\sigma_B^2+\sigma^2}{(y+\sigma^2W)^2}}{\left(E\frac{1}{y+\sigma^2W}\right)^2} \geq \frac{1}{E\frac{1}{\sigma_B^2+\sigma^2}}\,. \tag{11}$$

For moderate p this approximation could be rather poor when y is small. If the distributions of W and σ were given, then the asymptotically optimal value of

$y = y_{\mathrm{opt}}$, can be found as the minimizer of $S(y)$. However, commonly $S(y)$ is flat around its minimum, and the statistical estimates $\tilde{x}$ with $y = y_{\mathrm{opt}}$ have larger variance than the procedures detailed here. Observe that when $W \equiv 1$, $y_{\mathrm{opt}} = \sigma_B^2$, in which case the lower bound in (11) is attained.

When $\sigma^2 \equiv \sigma_0^2, S(y)$ monotonically decreases to the value $\sigma_B^2 + \sigma_0^2$, so that in this case $y_{\mathrm{opt}} = \infty$, and $\bar{x}$ is asymptotically optimal. In general, for $y \to \infty$,

$$\frac{S'(y)}{2E(\sigma_B^2 + \sigma^2)} \sim \frac{\mathrm{Var}(\sigma^2)}{y^2} > 0\,,$$

so that $y_{\mathrm{opt}} < \infty$.

Also, provided that $E\sigma^{-4} < \infty$, $S'(0) < 0$, unless $W \equiv 1$. Therefore, in this setting the Graybill–Deal estimator, for which $y = 0$, cannot be optimal for non-degenerate distributions. Since the weights s_i^2 do not utilize observations from other studies, which also provide information on σ_i^2 via the common parameter μ, this estimator has known deficiencies (see for example, Hinkley, 1979).

4. Approximate variances of the weighted means

In this Section we accept $S(y)/p$ as an approximation to the variance of the weighted means statistics, for which y in (7) found from estimating equations converges in probability, $y \to y_0$, as $p \to \infty$. For example, under our assumptions, the solution y_{MP} of (9) for the Mandel–Paule procedure (or its modified version) converges to y_0 found from the equation

$$E\frac{(X-\mu)^2}{y_0 + S^2} = E\frac{\sigma_B^2 + \sigma^2}{y_0 + \sigma^2 W} = 1\,. \tag{12}$$

Since, by Jensen's inequality,

$$E\frac{\sigma_B^2 + \sigma^2}{y_0 + \sigma^2 W} \geq E\frac{\sigma_B^2 + \sigma^2}{y_0 + \sigma^2 EW} = E\frac{\sigma_B^2 + \sigma^2}{y_0 + \sigma^2}\,,$$

the solution y_0 of (12) is larger than σ_B^2, unless $W \equiv 1$. It follows that the Mandel–Paule estimate y_0 of σ_B^2 typically is not consistent in our setting. An extensive simulation study of various (interval) estimators of σ_B^2 was performed by Wimmer and Witkovský (2002).

In our model for any fixed y the variance of $\tilde{x}$ can be estimated by

$$\delta_0 = \frac{p}{p-1}\sum_1^p \frac{(x_i - \tilde{x})^2}{(y + s_i^2)^2}\left[\sum_1^p \frac{1}{y + s_i^2}\right]^{-2}. \tag{13}$$

Indeed, as indicated above, $p^{-1}\sum_1^p \frac{1}{y+s_i^2} \to E\frac{1}{y+S^2}$, and

$$\frac{1}{p-1}\sum_1^p \frac{(x_i-\tilde{x})^2}{(y+s_i^2)^2} \approx E\frac{(X-\mu)^2}{(y+S^2)^2},$$

where the factor $(p-1)^{-1}$ gives a better approximation than the p^{-1} used by Rukhin and Vangel (1998), as

$$\sum_1^p \frac{x_i-\tilde{x}}{y+s_i^2} = 0.$$

Simulations show that $p\delta_0$, which is a consistent estimator of the variance given in (11), is somewhat better than traditional estimates

$$\delta_1 = \left[\sum_1^p \frac{1}{y+s_i^2}\right]^{-1}$$

unless σ_B^2 (and y) are very small, as then δ_0 underestimates the variance of $\tilde{x}$. Observe that δ_1 cannot always give a good estimate of the variance of the weighted means statistic with the weights (7), as this would suggest that the minimal value of the variance be attained at $y=0$.

Therefore, if z_α denotes the critical point of the standard normal distribution, the interval,

$$\tilde{x} \pm z_{\alpha/2}\frac{\sqrt{p\sum_1^p \frac{(x_i-\tilde{x})^2}{(y+s_i^2)^2}}}{\sqrt{p-1}\sum_1^p \frac{1}{y+s_i^2}}, \tag{14}$$

is an approximate $(1-\alpha)$%-confidence interval for the treatment effect μ on the basis of the weighted means statistics $\tilde{x}$. Notice that this interval does not depend on the specific form of the distribution of W or σ^2.

For sufficiently large p it is more reasonable to estimate the variance of the Mandel–Paule rule $\tilde{x}$ using $\delta_0 = \delta_{0,\mathrm{MP}}$, with $y = y_{\mathrm{MP}}$ determined by (9), rather than by

$$\delta_{1,\mathrm{MP}} = \left[\sum_1^p \frac{1}{y_{\mathrm{MP}}+s_i^2}\right]^{-1}, \tag{15}$$

as advocated in Mandel (1991) p. 72. Indeed, the limit of $p\delta_1$ is not (11), but $\left[E\frac{1}{y_0+S^2}\right]^{-1}$, so that this estimator of the limiting variance of the Mandel–Paule procedure may not even be consistent. Mandel (1991) himself writes:

$(y_{\mathrm{MP}} + s_i^2)^{-1}$ "are actually only sample estimates of the true weights resulting in perhaps considerable uncertainty in" δ_1.

For large p, δ_1 underestimates the limiting variance (11) of the Mandel–Paule rule. To see that, it suffices to establish the inequality

$$\frac{E\frac{(X-\mu)^2}{(y_0+S^2)^2}}{\left(E\frac{1}{y_0+S^2}\right)^2} \geq \frac{1}{E\frac{1}{y_0+S^2}} = \frac{E\frac{(X-\mu)^2}{y_0+S^2}}{E\frac{1}{y_0+S^2}},$$

which means that

$$E\frac{\sigma_B^2+\sigma^2}{(y_0+\sigma^2 W)^2} \geq \left(E\frac{\sigma_B^2+\sigma^2}{y_0+\sigma^2 W}\right)\left(E\frac{1}{y_0+\sigma^2 W}\right).$$

This inequality follows from the fact that for any fixed value of σ two functions of W, $(\sigma_B^2+\sigma^2)/(y_0+\sigma^2 W)$ and $1/(y_0+\sigma^2 W)$, are positively correlated, so that

$$E\left(\frac{\sigma_B^2+\sigma^2}{y_0+\sigma^2 W}\left[\frac{1}{y_0+\sigma^2 EW} - E\frac{1}{y_0+\sigma^2 EW}\right]\Bigg|\,\sigma\right) \geq 0.$$

Thus, the formula (15) tends to underestimate the true value of the variance when p is large.

A similar situation holds for the DerSimonian–Laird procedure as the value y_{DL} from (8) converges to

$$y_0 = \max\left[0, \frac{E\frac{(X-\mu)^2}{S^2} - 1}{E\frac{1}{S^2}}\right] = \sigma_B^2 + \frac{E\frac{1}{W} - 1}{E\frac{1}{\sigma^2}E\frac{1}{W}}. \tag{16}$$

Since $EW^{-1} \geq (EW)^{-1} = 1$, this y_0 also overestimates σ_B^2, unless $W \equiv 1$.

For sufficiently large p, the variance of the DerSimonian–Laird estimator is better estimated by

$$\delta_{0\,\mathrm{DS}} = \sum_1^p \frac{(x_i - \tilde{x})^2}{(y_{\mathrm{DL}} + s_i^2)^2}\left[\sum_1^p \frac{1}{y_{\mathrm{DL}} + s_i^2}\right]^{-2}, \tag{17}$$

than by the commonly used rule,

$$\delta_{1\,\mathrm{DS}} = \left[\sum_1^p \frac{1}{y_{\mathrm{DL}} + s_i^2}\right]^{-1}$$

mentioned in Section 2.

Notice that for the ANOVA type estimate of σ_B^2, the corresponding value y converges to

$$y_0 = \max\left[0,\, E(\sigma^2 + \sigma_B^2) - E\sigma^2\right] = \sigma_B^2,$$

so this method leads to a consistent estimator.

Theorem 4.1. *Under conditions of Section 3 for any weighted means statistic* (3), $\sqrt{p}[\tilde{x} - \mu]$ *is approximately normal with zero mean and the variance given by* (11). *For the Mandel–Paule procedure the limiting value* y_0 *is found from* (12), *and this procedure cannot be a consistent estimate of* σ_B^2. *For the DerSimonian–Laird rule the value of* y_0 *is determined from* (16). *The interval* (14) *is an approximately* $(1-\alpha)$ *%-confidence interval for* μ, *as the statistic* (13) *estimates the variance of the limiting distribution of* $\tilde{x}$, *consistently whereas* $\delta_{1\,\mathrm{MP}}$ *for the Mandel–Paule statistic and* $\delta_{1\,\mathrm{DS}}$ *for the DerSimonian–Laird rule are not consistent estimators of this variance.*

5. Simulation results

Some results of a Monte Carlo simulation study with $p = 10, 20, 30$, $\sigma_B^2 = 0, 1, 5$, $\nu = 5$, and G an inverted gamma-distribution with parameters $\alpha = 2$, $\beta = 1$, are shown in the Tables 1–3. Table 1 provides the mean squared errors of the corresponding estimators (8), (9) and (10). These mean squared errors are rescaled by p. Table 2 contains the values of y_{MP}, the Mandel–Paule choice, of y_{MMP}, the modified Mandel–Paule rule, and of y_{DL}, the DerSimonian–Laird method value. Table 3 reports the coverage probabilities of the confidence intervals $\tilde{x} \pm 2\sqrt{\delta_0}$ and $\tilde{x} \pm 2\sqrt{\delta_1}$ for the three estimators under study.

Table 1. Values of the Monte-Carlo mean squared errors (multiplied by p) when $\sigma_B^2 = 0, 1, 5$ and $p = 10, 20, 30$.

	MP	MMP	DL
$p = 10,\ \sigma_B^2 = 0$	0.71	0.73	0.69
$p = 20,\ \sigma_B^2 = 0$	0.67	0.68	0.65
$p = 30,\ \sigma_B^2 = 0$	0.69	0.70	0.68
$p = 10,\ \sigma_B^2 = 1$	1.81	1.83	1.82
$p = 20,\ \sigma_B^2 = 1$	1.79	1.80	1.79
$p = 30,\ \sigma_B^2 = 1$	1.75	1.75	1.75
$p = 10,\ \sigma_B^2 = 5$	6.07	6.08	6.08
$p = 20,\ \sigma_B^2 = 5$	5.84	5.84	5.85
$p = 30,\ \sigma_B^2 = 5$	5.85	5.84	5.89

Note that the original Mandel–Paule rule, whose mean squared error is almost the same as that of (8), exhibits somewhat better performance than the

Table 2. Values of y for three estimators.

	y_{MP}	y_{MMP}	y_{DL}
$p = 10,\ \sigma_B^2 = 0$	0.21	0.17	0.21
$p = 20,\ \sigma_B^2 = 0$	0.19	0.17	0.20
$p = 30,\ \sigma_B^2 = 0$	0.18	0.16	0.19
$p = 10,\ \sigma_B^2 = 1$	1.10	0.95	1.13
$p = 20,\ \sigma_B^2 = 1$	1.11	1.05	1.17
$p = 30,\ \sigma_B^2 = 1$	1.11	1.06	1.18
$p = 10,\ \sigma_B^2 = 5$	4.98	4.47	4.91
$p = 20,\ \sigma_B^2 = 5$	5.09	5.84	5.84
$p = 30,\ \sigma_B^2 = 5$	5.07	5.09	5.17

Table 3. Coverage probabilities for the confidence intervals $\tilde{x} \pm 2\sqrt{\delta_0}$ (upper line) and $\tilde{x} \pm 2\sqrt{\delta_1}$ (lower line) for the three estimators.

	δ_{MP}	δ_{MMP}	δ_{DL}
$p = 10,\ \sigma_B^2 = 0$	0.88 0.91	0.88 0.90	0.88 0.91
$p = 20,\ \sigma_B^2 = 0$	0.92 0.92	0.92 0.92	0.92 0.93
$p = 30,\ \sigma_B^2 = 0$	0.93 0.92	0.92 0.91	0.92 0.92
$p = 10,\ \sigma_B^2 = 1$	0.91 0.91	0.91 0.89	0.90 0.90
$p = 20,\ \sigma_B^2 = 1$	0.94 0.93	0.94 0.92	0.93 0.93
$p = 30,\ \sigma_B^2 = 1$	0.95 0.94	0.94 0.94	0.94 0.94
$p = 10,\ \sigma_B^2 = 5$	0.92 0.91	0.92 0.90	0.91 0.90
$p = 20,\ \sigma_B^2 = 5$	0.94 0.94	0.94 0.93	0.94 0.93
$p = 30,\ \sigma_B^2 = 5$	0.95 0.94	0.95 0.94	0.94 0.94

modified procedure (10). Also, in these simulations all three estimators above systematically outperform the estimator of the mean based on ANOVA MINQUE type estimators of σ_i^2 and σ_B^2 considered in Rao *et al.* (1981) by about 20 %.

Table 2 shows that the values of y_{MP}, y_{MMP} and y_{DL} have a tendency to stabilize for large values of p at a value exceeding σ_B^2. Also in these simulations the estimators δ_0 are indeed somewhat better than δ_1 for $p \geq 20$, but not for $p = 10$. The confidence coefficient of the interval (14) for all three estimators did not fall below 0.88 (for $\sigma_B^2 = 0$), while the nominal value was chosen to be about 95 %.

The evaluation of the full maximum likelihood estimator and its mean squared error is much more complicated. Some results concerning this rule are given in Vangel and Rukhin (1999).

In conclusion, we stress that the DerSimonian–Laird method is an excellent alternative to the Mandel–Paule algorithm, and, as such, has a great potential for interlaboratory studies. However, the corresponding y_{DS} or y_{MP} may not be appropriate estimates of the between-trials variance.

REFERENCES

[1] COCHRAN, W. G.: *Problems arising in the analysis of a series of similar experiments*, J. Roy. Statist. Soc., Supplement **4** (1937), 102–118.

[2] COCHRAN, W. G.: *The combination of estimates from different experiments*, Biometrics **10** (1954), 101–129.

[3] COX, D. R.: *A note on partially Bayes inference and the linear model*, Biometrika **62** (1975), 651–654.

[4] CROWDER, M.: *Interlaboratory comparisons: Round robins with random effects*, Appl. Statist. **41** (1992), 409–425.

[5] DERSIMONIAN, R—LAIRD, N.: *Meta-analysis in clinical trials*, Controlled Clinical Trials **7** (1986), 177–188.

[6] GRAYBILL, F. A.—DEAL, R. B.: *Combining unbiased estimators*, Biometrics **15** (1959), 543–550.

[7] HINKLEY, D. V.: *A note on the weighted means problem*, Scandinavian J. Statist. **6** (1979), 27–40.

[8] MANDEL, J.: *Evaluation and Control of Measurements*, Marcel Dekker, Inc., New York, 1991.

[9] MANDEL, J.—PAULE, R. C.: *Interlaboratory evaluation of a material with unequal number of replicates*, Analytical Chemistry **42** (1970), 1194–1197.

[10] PAULE, R. C.—MANDEL, J.: *Consensus values and weighting factors*, J. Res. Nat. Bureau Standards **87** (1982), 377–385.

[11] RAO, P. S. R. S.: *Cochran's contributions to variance component models for combining estimates*, in: W. G. Cochran's Impact on Statistics (P. Rao and J. Sedransk, eds.), J. Wiley, New York, 1981.

[12] RAO, P. S. R. S.—KAPLAN, J.—COCHRAN, W. G.: *Estimators for the one-way random effects model with unequal error variances*, J. Amer. Statist. Assoc. **76** (1981), 89–97.

[13] RUKHIN, A. L.—BIGGERSTAFF, B.—VANGEL, M. G.: *Restricted maximum likelihood estimation of a common mean and Mandel-Paule algorithm*, J. Statist. Plann. Inference **83** (2000), 319–330.

[14] RUKHIN, A. L.—VANGEL, M. G.: *Estimation of a common mean and weighted means statistics*, J. Amer. Statist. Assoc. **93** (1998), 303–308.

[15] SCHILLER, S.—EBERHARDT, K.: *Combining data from independent chemical analysis methods*, Spectrochimica Acta **12** (1992), 1607–1613.

[16] VANGEL, M. G.—RUKHIN, A. L.: *Maximum likelihood analysis for heteroscedastic one–way random effects ANOVA in interlaboratory studies*, Biometrics **55** (1999), 129–136.

[17] WIMMER, G.—WITKOVSKÝ, V.: *Between group variance component interval estimation for the unbalanced heteroscedastic one-way random effects model*, J. Statist. Comput. Simulation. **73** (2003), 333–346.

Received March 21, 2002

Department of Mathematics and Statistics
University of Maryland at Baltimore County
1000 Hilltop Circle
Baltilmore
MD–21250
U. S. A.

Statistical Engineering Division
National Institute of Standars and Technology
Building 820
Gaithersburg
MD–20899-0001
U. S. A

E-mail: rukhin@math.umbc.edu