

The applicability to systematic reviews of common effect, random effects and fixed effects approaches to meta-analysis

Statistical Methods in Medical Research

1–14

© The Author(s) 2026



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/09622802261439260

journals.sagepub.com/home/smmRichard J Stevens¹ 

Abstract

Systematic reviewers planning quantitative meta-analysis usually choose between fixed effect meta-analysis, and random effects meta-analysis. An alternative method is called fixed effects (note the s in the name). This method has the unique property that the target estimand is defined by the variances of the studies found by the systematic review. This article considers each in relation to the quantitative analysis of data to be obtained by systematic review. For clarity, I refer to the traditional fixed effect method as the common effect method and the newer approach as fixed effects (plural). Case studies illustrate the *post hoc* nature of the fixed effects (plural) method, in which the population under study is determined by the data rather than by the protocol. Mathematical analysis shows that unlike common effect and random effects methods, the fixed effects (plural) method requires an additional, and unrealistic, assumption about the data obtained in systematic reviews. A simulation study demonstrates that confidence intervals from fixed effects (plural) meta-analysis do not account for the *post hoc* nature of the method. Fixed effects (plural) meta-analysis is neither a slot-in replacement for the common effect method nor for the random effects method of meta-analysis. Of the three methods considered here, the common effect method and the random effects method are potentially valid for the quantitative analysis of systematic reviews.

Keywords

Meta-analysis, systematic review, common effect meta-analysis, random effects meta-analysis, fixed effects meta-analysis, ancillarity

1 Background

Many of us were taught, and have taught, that systematic reviewers planning a quantitative meta-analysis typically choose between two modelling approaches: on the one hand, methods based on a fixed effect model, on the other hand, methods based on a random effects model. The fixed effect model makes a modelling assumption that the true value of the effect of interest is the same in all included studies eligible for inclusion in the meta-analysis. The random effects model explicitly assumes a variety of different true values are possible, and models their distribution.¹ However, it is no longer the only model that allows for multiple true values. Rice et al.² introduced a third approach to meta-analysis, which has computational similarities to fixed effect meta-analysis even while using a model that is more general than random effects meta-analysis. Rice et al. call their model by the name ‘fixed effects’, for which reason some authors in recent years have begun to find new names for the long-established fixed effect method. In this article, I will refer to the long-established fixed effect approach as ‘common effect’ meta-analysis, and the approach of Rice et al. as ‘fixed effects (plural)’ meta-analysis, for clarity.

¹Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK

Corresponding author:

Richard J Stevens, Nuffield Department of Primary Care Health Sciences, University of Oxford, New Radcliffe House, Radcliffe Primary Care Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford OX2 6GG, UK.

Email: richard.stevens@phc.ox.ac.uk

In an era of protocol-driven medical research, a systematic literature review begins with a protocol that prespecifies the aims and methods. The research question should be clearly framed.³ Structured formats for framing a research question about treatment include the PICO format, in which *P* denotes that the population of interest should be specified; *I* and *C* denote that the intervention and comparator should be specified, respectively; and *O* denotes that the outcome should be specified.⁴ When a systematic review is to be analysed quantitatively, the analysis methods should be well-chosen for the research question. In this article, I explore the applicability to systematic reviews of each of the three approaches to meta-analysis discussed above. A primary objective for the article is to examine whether the fixed effects (plural) approach can be considered simply a ‘re-evaluation’ of the common effect (formerly, ‘fixed effect’) approach, or whether it needs to be considered something wholly new. A secondary objective is to examine whether the fixed effects (plural) approach can be considered the most general of the three.

In Section 2, I establish notation for this article and use it to state formally the model assumptions, estimands and estimates for each approach. In Section 4.1, I illustrate the substantial differences in interpretation by applying each to a small case study. In Section 3, I show that an additional assumption is required for the fixed effects (plural) method. In Section 5, I conduct a simulation study of the performance of each approach as heterogeneity increases.

Throughout this article, I consider estimation of a single effect parameter. Quantification and testing of heterogeneity, and other developments such as bivariate meta-analysis and network meta-analysis, are beyond the scope here. It will be useful to distinguish between models and methods. In particular, there are a great many methods to estimate the parameters of the random effects model.⁵ By ‘model’ I mean a set of assumptions expressed mathematically. In ‘method’ I include the choice of estimand as well as the formulae used for point, variance and interval estimation about an estimand of interest.

2 Models and methods

2.1 Common effect model and method

Suppose k studies yield scalar effect estimates ($X_i; i = 1, \dots, k$) with standard errors (or estimated standard errors) ($\sigma_i; i = 1, \dots, k$). For a simple common effect model, we can write

$$X_i \sim \mathcal{N}(\mu, \sigma_i^2) \quad (1)$$

where ($X_i; i = 1, \dots, k$) are k effect estimates from k different studies. For the moment, we assume also that the σ_i^2 are fixed constants, but we return to this point in Section 3. It is also common to assume that the σ_i^2 are known exactly; in practice of course, only estimates of σ_i^2 are available, but it has been shown that this has little impact on results.⁶

In this approach, the target estimand is a parameter μ that is assumed to be common to all the populations in which the studies yielding the $X_i, i = 1, \dots, k$ have been carried out. The usual estimate of μ is

$$\hat{\mu} = \frac{\sum_{i=1}^k X_i \sigma_i^{-2}}{\sum_{i=1}^k \sigma_i^{-2}} \quad (2)$$

This has been called the precision-weighted average, because the σ_i^{-2} are referred to as the precisions of the studies, or the inverse-variance weighted average. The variance estimate

$$\text{Var}(\hat{\mu}) = \left(\sum_{i=1}^k \sigma_i^{-2} \right)^{-1}$$

can be used to construct interval estimates for μ .

Strictly speaking, equation (2) should use the estimated variances $\hat{\sigma}_i^2, i = 1, \dots, k$ rather than the true variances. The distinction is not important unless the studies are small⁶ and so for simplicity I omit the distinction throughout this section, but return to it briefly in Section 3.

Other estimation methods exist for the common effect model.^{7,8} These are specific to the case that the $\hat{\mu}$ are effect measures, such as (log) odds ratios or (log) risk ratios, for a binary outcome, and that the exposure of interest is also binary. The method (2), on the other hand, is generic and can be used for analysis of both binary and continuous data.

Typical methods based on the common effect model have the attractive property that study weights are proportional to the information size of the study.

Note on terminology: the approach described here as common effect meta-analysis is that originally known as fixed effect meta-analysis; the name ‘common effect’ is increasingly used to avoid confusion with the fixed effects approach described in Section 2.3.

2.2 Random effects model and methods

In the random effects approach, we no longer assume a single constant μ . Instead, we assume that in each of the studies $i = 1, \dots, k$ in which the X_i have been estimated, a different μ_i exists; and further, that the μ_i are distributed Normally around some central μ_R , as follows:

$$X_i \sim \mathcal{N}(\mu_i, \sigma_i^2), \text{ where} \quad (3)$$

$$\mu_i \sim \mathcal{N}(\mu_R, \tau^2) \quad (4)$$

Here μ_R is a central effect parameter and τ^2 is the variance of the study-specific effect parameters. For some purposes, it is possible to relax the normal distribution assumption, and assume only that the $\mu_i, i = 1, \dots, k$ are sampled from a distribution with expectation μ and variance τ^2 .⁹ Again, we usually assume the σ_i^2 to be known constants. The impact of assuming them to be known, rather than estimated, has been studied elsewhere. In Section 3, we return to the assumption that they are constants.

It is common to use

$$\hat{\mu}_R = \frac{\sum_{i=1}^k X_i (\sigma_i^2 + \hat{\tau}^2)^{-1}}{\sum_{i=1}^k (\sigma_i^2 + \hat{\tau}^2)^{-1}} \quad (5)$$

as a point estimate of μ_R . The target estimand here is now the central μ_R of the hypothesised normal distribution, so the interpretation as well as the method of estimation differs from the common effect approach.

In practice, τ is not known. Many different estimates of τ have been proposed, and also multiple approaches to interval estimation. There are therefore many different random effects methods, for a single random effects model. In the random effects model, τ^2 is not merely a nuisance parameter in the study of μ_R , but may be of importance in its own right. The random effects model, therefore, differs from the common effect model in having two rather than one potential estimands of interest.

Methods based on the random effects model have the attractive property that numerical heterogeneity between studies is reflected in the interval estimate – the more studies disagree, the wider the confidence interval. This is not true of common effect methods, in which confidence interval calculation is oblivious to the numerical homogeneity or heterogeneity between studies. However, random effects methods do not necessarily give studies weights that well-reflect the information size of the study, especially when heterogeneity is present: the smallest studies may have much more influence on the summary estimate than they would in a common effect meta-analysis.

2.3 Fixed effects (plural) model and method

Rice et al. propose a model in which the $\mu_i, i = 1, \dots, k$ are study-specific means, as in the random effects model (3), but (in contrast to the random effects model) there is no assumption made about their distribution. The terminology ‘fixed effects’, where ‘effects’ is plural, is intended to suggest that there are multiple true μ_i but that no random distribution is assumed for them. If we write this model as follows:

$$X_i \sim \mathcal{N}(\mu_i, \sigma_i^2) \quad (6)$$

while taking the $\mu_i, i = 1, \dots, k$ to be unknown parameters, then it is a generalisation of the random effects model, in the sense that it is always true when (3) is true, but not vice versa. An alternative interpretation is presented in Section 6.3

As in the previous sections, we assume first that the σ_i^2 are constants, and that they are known exactly. Dominguez-Islas and Rice⁶ show that the assumption that they are known exactly can be relaxed. We return later (Section 3) to the assumption that they are constants.

Rice et al. then propose that the following estimand be of interest: ($\mu_i; i = 1, \dots, k$):

$$\beta_F = \frac{\sum_{i=1}^k \mu_i \sigma_i^{-2}}{\sum_{i=1}^k \sigma_i^{-2}} \quad (7)$$

The motivation for this estimand is that, if we take this to be the estimand of interest, and make the constant variance assumption, then (2) is the least squares estimate. In this way, Rice et al. propose that the same estimation method can be used for this ‘fixed effects (plural)’ model as is currently used for the common effect model. An additional estimand, to quantify heterogeneity, is proposed in a companion paper.⁶

The fixed effects (plural) approach, therefore, differs from the common effect approach and the random effects approach in that the estimand itself is a combination of location measures and variances. In the traditional common effect method, variances appear in the formula for the estimate, but the thing we have decided to estimate, the *estimand*, is simply μ . In the fixed effects (plural) method, variances also appear in the definition of the target estimand.

3 Examination of the constant variance assumption

Arguably, the sizes, variances and weights of studies in a meta-analysis should be considered random variables.¹⁰ In the exposition in Sections 2.1 to 2.3, we made two assumptions about the study variances σ_i^2 : firstly, that they are constants, and secondly, that they are known exactly. Since the second assumption has been studied previously, in this article, I consider the assumption that the σ_i^2 can be considered to be constants.

Conditioning on the study-specific variances, as if they were constants, may seem analogous to conditioning on the sample size n in an analysis of a single dataset. One defence for treating such an n as constant is that it is, usually, ancillary for the estimands of interest. Here, I examine whether such ancillarity arguments can be applied to the study-specific variances in the meta-analysis of systematic review data.

In this section, I outline the straightforward proofs (Sections 3.1 and 3.2) that in a common effect meta-analysis, or a random effects meta-analysis, the study-specific variances are ancillary for the effect parameters, in order to contrast this (Section 3.3) with the fixed effects (plural) meta-analysis model.

Noting that ‘ancillary’ has multiple definitions,¹¹ for this article, I use the following: call a ancillary for a parameter θ if the maximum likelihood estimate of θ is the same whether we treat a as a known constant or as random. The definition of ‘ancillary’ given in Section 7.3 of Pawitan¹² is sufficient for this.

3.1 Common effect meta-analysis

We can show as follows that for common effect meta-analysis, estimation of the parameter μ is agnostic to whether the true within-study variances are assumed known or assumed random.

If we consider the $\sigma_i, i = 1, \dots, k$ to be known constants, then the likelihood function of μ given the data $X_i, i = 1, \dots, k$ is

$$L(\mu; X_i, i = 1, \dots, k) = \frac{1}{\prod \sigma_i} \times \prod \phi\left(\frac{\mu - X_i}{\sigma_i}\right)$$

where \prod denotes product from $i = 1$ to k and $\phi(\dots)$ denotes the probability density function (pdf) of the standard normal distribution. Thus, the log-likelihood of μ given $X_i, i = 1, \dots, k$ is

$$l(\mu; X_i, i = 1, \dots, k) = \text{constant} - \frac{1}{2} \sum_{i=1}^k \frac{(\mu - X_i)^2}{\sigma_i^2}$$

The first derivative of $l(\mu)$ with respect to μ is then

$$\frac{\partial l}{\partial \mu} = - \sum_{i=1}^k \frac{(\mu - X_i)}{\sigma_i^2}$$

This could, for example, be used to derive expression (2) as the maximum likelihood estimate of μ , if maximum likelihood estimation were thought appropriate.

Suppose instead we choose to regard the $\sigma_i, i = 1, \dots, k$ as arising at random from some distribution that has probability density function $f_\sigma(\sigma_i; \lambda)$. We will write

$$\sigma_i \sim f_\sigma(\sigma_i; \lambda) \tag{8}$$

Now the likelihood is

$$L(\mu, \lambda; X_i, \sigma_i, i = 1, \dots, k) = \prod f_\sigma(\sigma_i; \lambda) \times \frac{1}{\prod \sigma_i} \times \prod \phi\left(\frac{\mu - X_i}{\sigma_i}\right)$$

and the log-likelihood of μ and λ given the data $X_i, \sigma_i, i = 1, \dots, k$ is

$$l(\mu, \lambda; X_i, \sigma_i, i = 1, \dots, k) = \sum_{i=1}^k \log f_{\sigma}(\sigma_i; \lambda) - \sum_{i=1}^k \log \sigma_i + \text{constant} - \frac{1}{2} \sum_{i=1}^k \frac{(\mu - X_i)^2}{\sigma_i^2} \quad (9)$$

Differentiating $l(\mu, \lambda)$ with respect to μ , the first three terms disappear, and the first derivative of $l(\mu)$ with respect to μ is again

$$\frac{\partial l}{\partial \mu} = - \sum_{i=1}^k \frac{(\mu - X_i)}{\sigma_i^2}$$

We have shown that the variances $\sigma_i^2, i = 1, \dots, k$ are ancillary for estimation of μ . In this model, inference about μ does not require an assumption that the study-specific variances are known constants.

3.2 Random effects meta-analysis

The arguments for the common effect model also extend to a random effects model for meta-analysis, such as (3).

If we consider the $\sigma_i, i = 1, \dots, k$ to be known constants, then the likelihood function of μ_R, τ given the data $X_i, i = 1, \dots, k$ is

$$L(\mu_R, \tau; X_i, i = 1, \dots, k) = \frac{1}{\prod \sqrt{\sigma_i^2 + \tau^2}} \times \prod \phi \left(\frac{\mu_R - X_i}{\sqrt{\sigma_i^2 + \tau^2}} \right)$$

where, as before, \prod denotes product from $i = 1$ to k . Thus, the log-likelihood of μ_R, τ given $X_i, i = 1, \dots, k$ is

$$l(\mu_R, \tau; X_i, i = 1, \dots, k) = \text{constant} - \sum_{i=1}^k \log(\sigma_i^2 + \tau^2) - \frac{1}{2} \sum_{i=1}^k \frac{(\mu_R - X_i)^2}{\sigma_i^2 + \tau^2}$$

Differentiating $l(\mu_R, \tau)$ with respect to μ_R and solving for 0, the first two terms disappear and the first derivative of $l(\mu_R)$ with respect to μ_R is

$$\frac{\partial l}{\partial \mu_R} = - \sum_{i=1}^k \frac{(\mu_R - X_i)}{\sigma_i^2 + \tau^2}$$

Suppose instead the $\sigma_i, i = 1, \dots, k$ are observations from a random distribution

$$\sigma_i \sim f_{\sigma}(\sigma_i; \lambda)$$

Now the log-likelihood of μ_R, τ given observations $X_i, \sigma_i, i = 1, \dots, k$ is

$$l(\mu_R, \lambda; X_i, \sigma_i, i = 1, \dots, k) = \sum_{i=1}^k \log f_{\sigma}(\sigma_i; \lambda) + \text{constant} - \sum_{i=1}^k \log(\sigma_i^2 + \tau^2) - \frac{1}{2} \sum_{i=1}^k \frac{(\mu_R - X_i)^2}{\sigma_i^2 + \tau^2}$$

When differentiating with respect to μ_R , all but the last term disappear, and so

$$\frac{\partial l}{\partial \mu_R} = - \sum_{i=1}^k \frac{(\mu_R - X_i)}{\sigma_i^2 + \tau^2}$$

as before. As in the common effect model, the study-specific variances are ancillary for inference about μ_R . They are also ancillary for τ^2 , by similar logic. Thus it is not necessary, in the random effects model, to assume the $\sigma_i^2, i = 1, \dots, k$ are known constants.

3.3 Fixed effects (plural) meta-analysis

The arguments in Section 3.1 extend to estimation of the $\mu_i, i = 1, \dots, k$. It can be shown that for any of the μ_i , the maximum likelihood estimate is the same whether we conceive the $\sigma_i, i = 1, \dots, k$ to be known constants or to be realisations from some distribution with pdf $f_\sigma(\sigma_i; \lambda)$.

However, the fixed effects (plural) method is not proposed for estimation of any μ_i , but for estimation of a β_F that is calculated from the $\mu_i, i = 1, \dots, k$ and the $\sigma_i, i = 1, \dots, k$. We now consider this, first for the case that the $\sigma_i, i = 1, \dots, k$ are conceived to be known constants, and then for extensions in which they are realisations from a random distribution.

Case 1: Within-study variances treated as constants. First, take the $\sigma_i, i = 1, \dots, k$ to be known constants. The model (6) then has k parameters, ($\mu_i; i = 1, \dots, k$).

Before we can consider a likelihood function for estimation of Rice et al.'s β_F we must write β_F as a parameter. This can be done by defining a re-parameterisation of the model as follows:

$$\begin{aligned}\beta_1 &= \sum_{i=1}^k w_i \mu_i \\ \beta_i &= \mu_i, \quad \forall i > 1\end{aligned}$$

where $w_i = \sigma_i^{-2} / \sum_{j=1}^k \sigma_j^{-2}$ are the usual inverse variance weights.

We can see that this is a valid re-parameterisation in that the model above for the probability distribution of the data $X_i, i = 1, \dots, k$ can be written in terms of the k parameters $\beta_i, i = 1, \dots, k$:

$$\begin{aligned}X_1 &\sim \mathcal{N}\left(\frac{\beta_1}{w_1} - \frac{1}{w_1} \sum_{j=2}^k w_j \beta_j, \sigma_1^2\right) \\ X_i &\sim \mathcal{N}(\beta_i, \sigma_i^2), \quad \forall i > 1\end{aligned}\tag{10}$$

Under the assumption that the weights $w_i, i = 1, \dots, k$ are constants, this is a valid reparameterisation, with data on the left hand side of each relation and parameters and constants on the right hand side. In this parameterisation, β_1 is the β_F of Rice et al. (equation (3) of that paper; equation (7) above). Thus, when the study-specific variances are constants, the β_F of Rice et al. is a well-defined parameter in the sense that it is one component of a parameter vector that can be used to specify the model (6).

The log-likelihood is then

$$l(\beta; X) = \text{constant} - \frac{1}{2} \left(\frac{\frac{\beta_1}{w_1} - \frac{1}{w_1} \sum_{j=2}^k w_j \beta_j - X_1}{\sigma_1} \right)^2 - \frac{1}{2} \sum_{i=2}^k \left(\frac{\beta_i - X_i}{\sigma_i} \right)^2$$

Let us take partial derivatives with respect to β_1 , the parameter of interest. Since we are assuming the σ_i^2 s and hence the w_i s to be constants, it follows that

$$\begin{aligned}\frac{\partial l}{\partial \beta_1} &= -\frac{1}{w_1 \sigma_1} \left(\frac{\frac{\beta_1}{w_1} - \frac{1}{w_1} \sum_{j=2}^k w_j \beta_j - X_1}{\sigma_1} \right) \\ &\propto \frac{\beta_1}{w_1} - \frac{1}{w_1} \sum_{j=2}^k w_j \beta_j - X_1\end{aligned}$$

This could be used to show that, under the assumption that the variances σ_i^2 are constants, the maximum likelihood estimate of Rice et al.'s β_F is the precision-weighted mean (2).

In summary, in the case that the variances are conceived to be known constants, β_F is a parameter and the maximum likelihood estimate of β_F is the precision weighted mean (2). The proofs in Dominguez-Islas and Rice⁶ show that this can be extended to the case that the variances are not known but estimated.

Extension to case 1: Within-study variances are constants, but must be estimated. Dominguez-Islas and Rice⁶ examine the distinction between the true study variances $\sigma_i^2, i = 1, \dots, k$, which are assumed to be unknown constants, and the estimated variances $s_i^2, i = 1, \dots, k$, assuming that each study $i = 1, \dots, k$ is ‘large enough’ for $\dots (\sigma_i^2)$ to be approximated with negligible error by its estimate s_i^2 . Then the estimand β_F is defined by the true weights, $w_i \propto \sigma_i^{-2}, i = 1, \dots, k$, while the estimator $\hat{\beta}_F$ is defined by the observed weights, $\hat{w}_i \propto \hat{\sigma}_i^{-2}, i = 1, \dots, k$. We can extend the arguments above to this scenario. The reparameterisation (10) uses the true weights, w_i , and true variances, σ_i^2 , and the argument holds.

Case 2: Within-study variances treated as random variables. Now suppose we regard the $\sigma_i, i = 1, \dots, k$ as arising at random from some distribution that has probability density function $f_\sigma(\sigma_i; \lambda)$, as in (8). We then have a model with $k + 1$ parameters.

If we seek to replicate the arguments of the previous case, difficulties now arise. The weighted sum

$$\frac{\sum_{i=1}^k \mu_i \sigma_i^{-2}}{\sum_{i=1}^k \sigma_i^{-2}} \quad (11)$$

contains both parameters and random variables. As a result, formula (10) does not define a valid re-parameterisation. Thus, the claim that β_F is a valid parameter cannot be made if the variances $\sigma_i, i = 1, \dots, k$ are taken to arise at random.

Since we have already noted that $\sigma_i, i = 1, \dots, k$ is ancillary for the k parameters $\mu_i, i = 1, \dots, k$ in (6), it is tempting to formulate an argument using the invariance properties of likelihood methods. However, the invariance properties of likelihoods apply only to data-independent transformations of parameters. In the formulation considered here, β depends on data $\sigma_i, i = 1, \dots, k$.

If, as in Section 2.3, we take the fixed effects (plural) model to be a straightforward generalisation of the random effects model, it may seem surprising that a difficulty arises in the latter but not the former. To understand this, note that no difficulty would arise, in the fixed effects (plural) model, if we were interested in estimating any of the $\mu_i, i = 1, \dots, k$. It is the decision to target β_F for estimation that creates a challenge for the fixed effects (plural) method. For this reason, it is helpful to distinguish the fixed effects *model*, defined by (6), from the fixed effects *method*, which is defined by the decision to estimate β_F as well as by the choice of model. The fixed effects model is a generalisation of the random effects *model*; but the fixed effects *method* is *not* a generalisation of the random effects *method*.

Extension to case 2: Within-study variances treated as random variables that are not observed directly and must be estimated. Again, we may extend the argument to the case that the $\sigma_i^2, i = 1, \dots, k$ must be estimated. We now have a hierarchical model for the variances, in which the σ_i^2 arise from some distribution f_σ , and the s_i^2 in turn are inexact estimates of the σ_i^2 . In the fixed effects (plural) approach, the $\sigma_i^2, i = 1, \dots, k$ define the estimand β_F and the $s_i^2, i = 1, \dots, k$ define the estimator (2). It remains true that the weighted sum (11) contains random variables, σ_i^{-2} , as well as parameters, and therefore does not provide a reparameterisation. If we try to construct a weighted sum in terms of the parameters of f_σ , for example by substituting $E[\sigma_i]^{-2}$ for σ_i^{-2} , then it will reduce to

$$\beta_1 = \frac{1}{k} \sum_{i=1}^k \mu_i \quad (12)$$

and we have a reparameterisation, but $\beta_1 \neq \beta_F$. This is because, in Case 2 and its extension, the k studies are of equal importance *a priori*, although differing amounts of information may arise in the sampling process; whereas in Case 1 and its extension, the variances differ by design, reflecting some weighting scheme that we wish to target in defining our estimand β_F .

3.4 Alternatives to likelihood arguments

The above sections used a likelihood framework for convenience, but likelihood estimation is not necessarily optimal. The ancillarity arguments can also be made with other definitions of ancillarity: for example, it can be shown by integration that the marginal distribution of θ in the random effects model is free of the variances.

Papers by Domínguez Islas and Rice⁶ make the case for the fixed effects (plural) approach by a minimum variance (or, in a Bayesian framework, maximum information) argument.¹³ Minimum variance arguments are often made for using a particular *estimator* in statistics, but it is unusual to see a minimum variance argument being proposed to motivate an *estimand*. The proof of the minimum variance argument in Lemma 1 of Domínguez Islas and Rice⁶ requires that variances

be modelled as constants (whether or not they are known exactly) and, therefore, the fixed effects (plural) method still requires this additional assumption, which is not required for the common effect or random effects method.

3.5 Implications

Although the fixed effects (plural) *model* is more general than either the common effect model or the random effects model, the fixed effects (plural) *approach* is not. For β_F to be valid as the target estimand, an additional assumption is required.

Each method of estimation requires a particular assumption. The constant effect method requires the constant effect assumption, and the random effects method requires a distributional assumption about the effects. We have now seen that the fixed effects (plural) method requires that we are prepared to model the study-specific variances as constants. If we think of the study-specific variances as data that will arise during the review, then we cannot select this method: the model (6) may apply, but the estimand (7) does not. Previous work⁶ has shown that we can generalise from known constants to constants that are reported with error. The work here shows that we cannot generalise more widely to the case that they are not constants, but arise at random.

Investigators specifying either common effect meta-analysis or random effects meta-analysis need not debate whether the study-specific variances are data or constants: as shown here, this conceptual point does not affect inference for these two models. Only investigators specifying fixed effects (plural) meta-analysis need to adopt a position on whether the constant variance assumption is tenable.

4 Case studies

To examine the applicability of the different estimands to a systematic review, I apply common effect, random effects, and fixed effects (plural) meta-analysis to previously published systematic reviews. The common effect and fixed effects (plural) meta-analysis are applied by the inverse variance method, and random effects meta-analysis by the DerSimonian and Laird method.¹⁴ This is sufficient to illustrate the differing interpretations but is not meant to imply a preference for these methods in practice.

4.1 Minton, 2008: Methylphenidate for fatigue in cancer patients

For simplicity of exposition I first select a previously published meta-analysis of just two studies. This is intended to illustrate differences in interpretation; it is not intended to imply that it is wise to attempt random effects meta-analysis, in particular, with so few studies.

A systematic review by Minton et al.¹⁵ found two studies that compared methylphenidate to placebo for fatigue in cancer patients. The study of Bruera et al.¹⁶ recruited patients with any tumour type, not on chemotherapy. The study of Fleishman et al.¹⁷ recruited patients with any tumour type, on chemotherapy. Figure 1 of Minton et al. shows a random effects meta-analysis of the standardised mean difference of fatigue score.

Table 1 shows the results of analysing the same data with either a common effect, random effects, or fixed effects (plural) meta-analysis. For the primary analysis, heterogeneity is low and all three analyses yield the same numerical results. For the common effect method, these can be interpreted as a point and interval estimate of a hypothetical true effect that is hypothesised to be constant across the different populations and settings in the Bruera and Fleishman studies. For the random effects method, the results should be interpreted as a point and interval estimate of the centre of some distribution, hypothesised to be normal, of possible true effects in Fleishman, Bruera and hypothetical other studies. For the fixed effects (plural) method, the results should be interpreted as a point and interval estimate of the average effect in a cancer population that is 43% like that of the Bruera study and 57% like that of the Fleishman study.

The second half of Table 1 shows the results of analysing a secondary outcome, adverse events, for the same two studies. The common effect results should, again, be interpreted as a point and interval estimate of a true effect that is hypothesised to be constant across the different populations. That the weights are different (from those in the common effect analysis of the primary outcome) informs the estimates, but not the interpretation. The random effects results should, likewise, be interpreted as a point and interval estimate of the centre of a hypothetical Normal distribution of possible true effects. That the weights are different (from those in the random effects analysis of the primary outcome) informs the estimates, but not the interpretation. The fixed effects (plural) results should be interpreted as an estimated effect in a population that is 16% akin to the population of the Bruera study and 84% akin to the population of the Fleishman study. The weights are different from the corresponding analysis of the primary outcome, and these weights not only inform the estimates, but also define the target estimand and hence inform the interpretation.

Table 1. Estimates, weights and interpretations when applying three meta-analysis methods to two studies of medication in cancer patients.

Model	Point estimate (95% CI ^a)	Weight for Studies 1 and 2	Interpretation
Primary analysis: fatigue score (standardised mean difference)			
Common effect	-0.30(-0.54, 0.05)	43% and 57%	Estimated common effect.
Random effects	-0.30(-0.54, 0.05)	43% and 57%	Estimated average effect.
Fixed effects (plural)	-0.30(-0.54, 0.05)	43% and 57%	Estimated effect in a population that is 43% akin to the population of Study 1 and 57% akin to the population of Study 2.
Secondary analysis: adverse events (log odds ratio)			
Common effect	1.10(0.13, 2.07)	16% and 84%	Estimated common effect.
Random effects	0.65(-1.38, 2.68)	37% and 63%	Estimated average effect.
Fixed effects (plural)	1.10(0.13, 2.07)	16% and 84%	Estimated effect in a population that is 16% akin to the population of Study 1 and 84% akin to the population of Study 2.

^aCI denotes confidence interval.

Table 2. Corticosteroids before preterm delivery and respiratory distress in babies, after Crowley et al. (1990).¹⁸

Study	Treatment group Events/n	Control group Events/n	Odds ratio (95% CI ^a)	Weight ^a
Liggins et al.	10/36	15/26	0.28 (0.10, 0.82)	20%
Taeusch et al.	1/3	4/6	0.25 (0.01, 4.73)	3%
Gamsu et al.	4/29	7/39	0.73 (0.19, 2.78)	13%
Collaborative group	6/10	7/16	1.93 (0.39, 9.60)	9%
Morales et al.	17/53	32/52	0.30 (0.13, 0.66)	35%
Papageorgiou et al.	2/5	11/12	0.06 (0.00, 0.92)	3%
Morrison et al.	6/36	11/28	0.31 (0.10, 0.99)	17%
Random effects estimate			0.38 (0.22, 0.67)	
Fixed effects (plural) estimate			0.37 (0.23, 0.60)	

^aCI denotes confidence interval. Weights shown are from the fixed effects (plural) analysis.

For the common effect method, the primary analysis and secondary analysis are both applicable to the same population. When using the random effects method, the primary analysis and secondary analysis are both applicable to the same population, albeit a different population to the common effect analyses. The fixed effects (plural) method has the remarkable property that results from the primary analysis, and results from the secondary analysis, should not be considered applicable to the same population as each other.

4.2 Crowley, 1990: Corticosteroids before preterm delivery

Crowley et al.¹⁸ reviewed the literature on maternal corticosteroids before preterm delivery, including a meta-analysis of neonatal respiratory distress in seven trials. Table 2 shows the results of re-analysing the same data with random effects meta-analysis or fixed effects (plural) meta-analysis. Common effect meta-analysis yields the same numerical results as fixed effects (plural) meta-analysis and is not shown separately.

The common effect estimate is that the odds ratio for treated versus control children is 0.37, with 95% confidence interval from 0.23 to 0.60, in all relevant populations. The random effects estimate is that the average odds ratio for treated versus control children is 0.38, with 95% confidence interval from 0.22 to 0.67. The interpretation is that this is the average of different odds ratios that would apply in different populations, and the calculation assumes that these hypothetical different odds ratios are Normally distributed. As is often the case, there are too few studies to usefully test the normal distribution assumption.

The fixed effects (plural) meta-analysis estimates a treatment effect in a meta-population that is 35% like the population studied by Morales et al., 20% like that studied by Liggins et al., 17% like that studied by Morrison et al., and so on. This composition of the meta-population is defined by the information sizes of the seven studies that were found in the systematic review. To this extent, the analysis has addressed a research question that was specified *post hoc* by the data obtained rather than by the investigators *a priori*.

The systematic review has been updated and re-analysed many times since the 1990 publication by Crowley,¹⁹ each time yielding a revised estimate of the odds ratio. In the common effect (or ‘fixed-effect’, singular) framework used by the authors, these are revised estimates of the same population parameter. If the authors used a fixed effects (plural) framework, then each new estimate would be interpretable with respect to a different estimand relevant to a different population.

4.3 Implications

The point has been made before that, when interpreting a systematic review, the choice of a common effect or a random effects method ‘affects the interpretation of the summary estimates’.¹ Common effect, and random effects, meta-analysis address slightly different research questions. Fixed effects (plural) meta-analysis addresses a different research question from either, studying an estimand that applies to a conceptual ‘meta-population’, and an argument has been made that this approach, too, can be informative.⁶ The case studies illustrate that the ‘meta-population’ under study depends on the data obtained. Therefore, it will not be known at the time a systematic review protocol is written. Investigators considering this approach should be aware of this, and also that there is no guarantee that the same ‘meta-population’ will be under study across all primary and secondary analyses in the review. The Crowley example also illustrates the assumption, in common effect meta-analysis, that the population parameter is constant across different studies – in this instance, an odds ratio common to all relevant studies. With seven included studies, it also illustrates the difficulty of an untestable normal distribution assumption in many systematic reviews.

5 Simulation study

The argument is made in Section 3.3 that even if the FE (plural) model is a generalisation of the random effects *model*, estimates produced by the FE (plural) *method* are not comparable with the random effects *method*. A simulation study is presented here, with the aim of comparing the behaviour of confidence intervals under different methods.

The data generating method is a scheme that has been used by previous authors^{20,21} for the random effects model. I simulate values of $\tau^2 = 0, 0.05, 0.10$ and 0.15 , and $k = 5, 8, 10, 20$ or 35 . I take $\mu_R = 0$ throughout; when $\tau^2 = 0$, this also gives $\mu = 0$. The variance of each study in each meta-analysis is drawn from a χ_1^2 distribution, multiplied by 0.25 , and rejected and resampled if outside the interval $(0.009, 0.6)$. This model for simulating variances was originally proposed by Brockwell and Gordon²² for realism. In each case 160,000 simulations were conducted.

The estimands have been documented in Section 2. The methods of analysis were as follows. Common effect and fixed effects (plural) meta-analysis are conducted using the function `rma.uni` from the `metafor` package,²³ with option `method='FE'`. Since the common effect model is valid only when there is no heterogeneity, I omit to report results for this method when $\tau^2 > 0$. Since we are considering the fixed effects (plural) model to be more general than the random effects model, I show results for all values of τ^2 for this model. DerSimonian and Laird random effects meta-analysis¹⁴ is conducted with the same package with the option `method='DL'`. Also shown are results from random effects meta-analysis using restricted maximum likelihood (REML) estimation, using the option `method='REML'`, and using Hartung-Knapp/Sidik-Jonkman estimation and confidence intervals (options `method='SJ'`, `knha=TRUE`).^{5,24}

The measure of performance used was the proportion of confidence intervals that contain 0. For the common effect method, this can be interpreted as the coverage property of a confidence interval for a constant μ . For the random effects method, this can be interpreted as the coverage property of a confidence interval for a central μ_R . For the fixed effects (plural) method, this can **not** be interpreted as the coverage property, because the method is designed to estimate a different parameter; recall that the aim of this simulation is to illustrate that this distinction is of numerical as well as conceptual importance. A simulation to determine coverage for β_F would proceed differently (see Section 5.1).

Table 3 shows the percentage of 95% confidence intervals that contain 0 in a large number of simulated meta-analyses, in which the true effects have distribution $\mathcal{N}(0, \tau^2)$, for a variety of values of τ^2 . For the common effect method, 95% of intervals contain 0 when τ^2 is 0. These have an interpretation as coverage probabilities for confidence intervals for μ . Results for other values of τ^2 are not shown because the common effect model is not valid unless $\tau^2 = 0$. For the random effects model, the proportion of intervals containing 0 varies with the level of heterogeneity and the number of studies. These results have interpretation as coverage probabilities for μ_R . For the fixed effects (plural) method, when $\tau^2 = 0$ results are, by definition, the same as for the common effect method. As heterogeneity increases, fewer and fewer of the 95% intervals contain 0. However, these should not be interpreted as poor coverage, since the confidence intervals are designed for estimation of β_F , which is not necessarily 0. Thus this should *not* be interpreted as poor coverage of confidence intervals: rather, it is a quantitative illustration of the difference between inference about β_F and inference about the μ or μ_R .

Table 3. Percentage of confidence intervals, by three different meta-analysis methods, that contain 0, in 160,000 simulated meta-analyses in which the effects in each study have distribution $\mathcal{N}(0, \tau^2)$.

τ^2	k (studies)	Common effect	Random effects ^a			Fixed effects (plural)
			(DL)	(REML)	(HKSJ)	
0	5	95.0	96.4	96.2	95.8	95.0
0	8	95.0	96.3	96.1	96.5	95.0
0	10	95.0	96.3	96.1	96.7	95.0
0	20	95.0	96.3	96.0	97.1	95.0
0	35	94.9	96.0	95.8	97.2	94.9
0.05	5	–	89.5	89.3	94.2	77.8
0.05	8	–	90.1	89.9	94.8	75.8
0.05	10	–	90.5	90.4	95.0	75.1
0.05	20	–	92.1	92.0	95.5	73.6
0.05	35	–	93.2	93.2	95.7	72.9
0.1	5	–	87.6	87.3	93.9	67.5
0.1	8	–	89.1	88.9	94.4	64.7
0.1	10	–	90.0	89.8	94.6	63.6
0.1	20	–	92.2	92.2	95.1	61.4
0.1	35	–	93.5	93.5	95.3	60.5
0.15	5	–	86.8	86.6	93.8	60.9
0.15	8	–	88.9	88.7	94.3	57.4
0.15	10	–	89.9	89.8	94.5	56.4
0.15	20	–	92.5	92.6	95.0	53.9
0.15	35	–	93.7	93.8	95.1	52.8

^aDL, REML and HKSJ denote the DerSimonian and Laird, the restricted maximum likelihood, and the Hartung-Knapp/Sidik-Jonkman methods, respectively.

5.1 Implications

The common effect model is only applicable when it is tenable to assume homogeneity of true effect(s), as has been observed often before. The random effects model is more widely applicable; systematic reviewers should be aware that good coverage properties for confidence intervals can depend in part on the specific random effect methods chosen.^{25,26}

More interestingly, the results in Table 3 provide a quantitative demonstration that the fixed effects (plural) *method* is profoundly different, even if the *model* is a generalisation of both the others. The results in the last column show that inference about the estimand β_F must never be mistaken for inference about a central μ or μ_R . The reduced coverage for 0 does not show that the method is wrong: if we were to study coverage for β_F , it would be 95% as intended. Rather, it shows, and emphasizes, that β_F is a very different target estimand.

Objections could arise that this simulation is not valid for the fixed effects (plural) method, either because the data was simulated from the wrong model, or because the wrong results are reported. It cannot be that the simulation is not valid for the fixed effects (plural) *model*: if it is valid for the random effects model (3), then it is valid for (6) because the latter is more general. Regarding the reporting: the data presented is fit for this purpose, which is to demonstrate, quantitatively, differences between the methods. Only if the last column were mis-presented as coverage probabilities for the target estimand would the data being wrong.

I have declined to present results for estimation of β_F , because β_F is not a predefined population parameter. The ability of the precision-weighted average to estimate the precision-weighted estimand β_F could be considered an example of the ‘Texas sharp-shooter’ fallacy: it is easy to strike a target when we move the target according to where the strike lands. It is interesting to consider in what simulation study we would consider the coverage properties for a fixed β_F . Previous authors simulated the case that all studies in a systematic review are equally sized and the true effect estimates are evenly and symmetrically spaced around 0,⁶ which is less general than the simulation here. A simulation of a multi-centre study with centre sizes selected by design, as discussed in Section 6.3, could be of interest to trialists.

6 Discussion

6.1 Summary of findings

The fixed effects (plural) approach is different from others in that the estimand is not defined until the systematic review is complete, the studies identified and their numerical information extracted. This is in contrast to the claim of

Rice et al. (their Section 4, line 1) that the estimand is a ‘well-defined’ population parameter; instead, it is a data-dependent quantity. The case studies (Section 4) illustrate this, and the ancillarity arguments in Section 3 clarify how this arises. The mathematics also show that this is not the case for the common effect or random effects approach to meta-analysis.

The simulation study illustrates that these abstract points affect the applicability of the confidence intervals from a meta-analysis. The confidence intervals from a common effect meta-analysis or random effects meta-analysis are designed to have 95% coverage for a true population parameter. They take into account that if our systematic review had obtained different data, potentially from different studies, then we would have a different estimate of the same estimand (population parameter). The confidence interval from a fixed effects (plural) meta-analysis is different: it takes into account that if we had obtained different data from the same studies, we would have a different estimate; but it does not take into account that in this method, if we had obtained different data from different studies, we would have a different estimate of *a different estimand*, applicable to *a different population*. There are two levels of sampling in a meta-analysis, but fixed effects (plural) meta-analysis only acknowledges one.

To my knowledge, there is no other method in statistics in which the target of estimation is data-dependent. From a statistical perspective, confidence intervals in hierarchical data should reflect every level of sampling uncertainty. In evidence based medicine, the research question – including the population of interest – is expected to be specified in the protocol, not defined by the data collected. This is in part because we at least aspire to estimate benefits and harms with respect to the same population (it is true that in any systematic review, publication bias and reporting bias may make this difficult; this is usually regarded as a problem, or at least an undesirable limitation). In Section 4.1, we saw that the fixed effects (plural) method could not give us a consistent population of interest, even when analysing the same samples for different clinical outcomes.

6.2 Limitations

The arguments here are almost entirely concerned with estimation, omitting hypothesis testing for reasons of length. The simulation study is likewise limited, considering only a null central effect parameter and normal distributions; nevertheless it is sufficient to demonstrate the differences between approaches.

There are more approaches to meta-analysis than are considered here. There are advocates for unweighted means¹⁰ and for weighting by quality of paper.²⁷ There are random effects meta-analysis methods that make different distributional assumptions from (3).^{28,29} There are advocates for using the precision-weighted mean even while making a random effects assumption.²⁰ Some authors have sought to account for between-study heterogeneity without using the term ‘random effects’ to describe their approach.³⁰ In this article, I have restricted attention to three model-based approaches to meta-analysis, and the numerical findings in the case studies and the simulation study do not fully explore the wide variety of methods available for the random effects model in particular. The poor coverage properties of this random effects approach in some scenarios should not be over-interpreted: see elsewhere for better examinations of the coverage properties of random effects meta-analysis methods.^{25,26}

The use of likelihood arguments in Section 3 is not intended to imply a preference for likelihood estimation in meta-analysis. The finding that fixed effects (plural) meta-analysis requires a constant variance assumption is not restricted to a likelihood framework, as discussed in Section 3.4.

My treatment is primarily concerned with the estimation of effect measures. By extension, the results also have some relevance to testing hypotheses about these effect measures. These findings will be less relevant to investigators whose sole or primary interest is in quantifying heterogeneity.

6.3 Other applications of meta-analysis

The treatment here is intended to be applicable to the analysis of data obtained by systematic review. Another application of meta-analysis is the analysis of a multi-centre study. Here the $(\theta_i; i = 1, \dots, k)$ are unknown constants, rather than draws from an unknown distribution. Though these constants are unknown, they are determined by design in that the investigators have selected the study centres to reflect specific population strata of interest. The investigators also determine k itself, and the relative sizes of the study centres, by design. Then the estimand β_F is indeed a well-defined parameter. Assuming the investigators have chosen the (information) sizes of the centres to reflect the relative importance of each population strata to their research, the status of β_F as an estimand of interest can be defended. In this application, the remark in Section 2.3 that (6) is a generalisation of the random effects model does not hold. Further, in Section 3.3, Case 1 applies, and the difficulties in Case 2 do not arise.

Thus, the difficulties with applying fixed effects (plural) meta-analysis to systematic review data are specific to systematic reviews, and do not arise when analysing a multi-centre study by this method.

6.4 Relation to the literature

Consideration of the different estimands in the common effect model and random effects model for meta-analysis is far from new.¹ There is more novelty in the study of the fixed effects (plural) estimand. Its behaviour across outcome measures, seen in the case studies, is implicit in its definition but to my knowledge has not been made explicit before. The ancillarity of the study-specific variances for common meta-analysis and random effects meta-analysis may not have been formally demonstrated before, but could be expected by analogy with the ancillarity of sample size in primary studies. That they are not ancillary for the estimand of the fixed effects (plural) approach is, I believe, novel. The simulation study demonstrates that it has practical as well as theoretical importance. This is the main contribution of the simulation study, since the results for common effect and random effects meta-analysis are not novel.

6.5 Conclusions

Systematic review investigators considering fixed effects (plural) meta-analysis must be aware that it is not simply common effect meta-analysis with a new rationale. Therefore the two should not be confused, even though the common effect method was once known as ‘fixed effect’ meta-analysis, because the target of estimation is different. Nor is it a generalisation of random effects meta-analysis, again because the estimand is different, but also because fixed effects (plural) meta-analysis requires a new assumption. This is the only approach to meta-analysis in which systematic review investigators need to reflect on whether they consider the standard errors recorded on their data extraction sheets to be ‘data’ or ‘constants’.

Systematic review investigators planning a meta-analysis should consider the usefulness of the different estimands. The usefulness and interpretation of the estimand in random effects model has been much discussed, including in the many papers cited here. The applicability and interpretation of the fixed effects (plural) estimand, and the ‘meta-population’ it represents, are defined only *post hoc*, and this conceptual difficulty is reflected in its confidence intervals. It cannot be recommended for use in systematic review.

Acknowledgements

I am grateful to Jake Olivier of University of New South Wales, and Thomas Fanshawe, Rafael Perera, Jason Oke, and the medical statistics group at Nuffield Department of Primary Care Health Sciences, for comments that have assisted the direction of this paper. This research was funded by the National Institute for Health and Care Research (NIHR) Applied Research Collaboration Oxford and Thames Valley at Oxford Health NHS Foundation Trust. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

ORCID iD

Richard J Stevens  <https://orcid.org/0000-0002-9258-4060>

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

1. Riley RD, Higgins JPT and Deeks JJ. Interpretation of random effects meta-analyses. *BMJ* 2011; **342**: 964–967.
2. Rice K, Higgins JPT and Lumley T. A re-evaluation of fixed effect(s) meta-analysis. *J R Stat Soc Ser A: Stat Soc* 2018; **181**: 205–227.
3. Thomas J, Kneale D, McKenzie J, et al. Chapter 2: Determining the scope of the review and the questions it will address. In: Higgins JPT et al. (eds.) *Cochrane handbook for systematic reviews of interventions version 6.4*. Cochrane Collaboration, www.training.cochrane.org/handbook (2023, accessed 27 August 2024).
4. Strauss S, Glasziou P, Richardson S et al. *Evidence-based medicine: how to practice and teach it*. 4th ed. Oxford, UK: Elsevier, 2013.
5. Veroniki A, Jackson D, Viechtbauer W et al. Methods to estimate the between-study variance and its uncertainty in meta-analysis. *Res Synthesis Methods* 2016; **7**: 55–79.
6. Domínguez Islas C and Rice KM. Addressing the estimation of standard errors in fixed effects meta-analysis. *Stat Med* 2018; **37**: 1788–1809.

7. Yusuf S, Peto R, Lewis J, et al. Beta blockade during and after myocardial infarction: an overview of the randomized trials. *Prog Cardiovasc Dis* 1985; **27**: 335–371.
8. Greenland S and Robins J. Estimation of a common effect parameter from sparse follow-up data. *Biometrics* 1985; **41**: 55–68.
9. Jackson D and White IR. When should meta-analysis avoid making hidden normality assumptions? *Biom J* 2018; **60**: 1040–1058.
10. Shuster JJ. Empirical vs natural weighting in random effects meta-analysis. *Stat Med* 2010; **29**: 1259–1265.
11. Casella GBR. *Statistical inference/George Casella, Roger L. Berger*. Belmont: Duxbury Press, 1990. ISBN 0534119581.
12. Pawitan Y. *In all likelihood: statistical modelling and inference using likelihood*. Oxford: Clarendon Press, 2001. ISBN 9780198507659.
13. Domínguez Islas C and Rice KM. Bayesian approaches to fixed effects meta-analysis. *Res Synth Methods* 2022; **13**: 520–532.
14. DerSimonian R and Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; **7**: 177–188.
15. Minton O, Richardson A, Sharpe M et al. A systematic review and meta-analysis of the pharmacological treatment of cancer-related fatigue. *J Natl Cancer Inst* 2008; **100**: 1155–1166.
16. Bruera E, Valero V, Driver L, et al. Patient-controlled methylphenidate for cancer fatigue: a double-blind, randomized, placebo-controlled trial. *J Clin Oncol* 2006; **24**: 2073–2078.
17. Fleishman S, Lower E, Zeldis J et al. A phase II, randomized, placebo-controlled trial of the safety and efficacy of dexmethylphenidate (d-MPH) as a treatment for fatigue and “chemobrain” in adult cancer patients. *Breast Cancer Res Treat* 2005; **94**: S214.
18. Crowley P, Chalmers I and Keirse M. The effects of corticosteroid administration before preterm delivery: an overview of the evidence from controlled trials. *BJOG Int J Obstet Gynaecol* 1990; **97**: 11–25. Cited By 760.
19. McGoldrick E, Stewart F, Parker R et al. Antenatal corticosteroids for accelerating fetal lung maturation for women at risk of preterm birth. *Cochrane Database Syst Rev* 2020; **2021**: CD004454. Cited By 601.
20. Henmi M and Copas JB. Confidence intervals for random effects meta-analysis and robustness to publication bias. *Stat Med* 2010; **29**: 2969–2983.
21. Sidik K and Jonkman JN. A simple confidence interval for meta-analysis. *Stat Med* 2002; **21**: 3153–3159.
22. Brockwell SE and Gordon IR. A comparison of statistical methods for meta-analysis. *Stat Med* 2001; **20**: 825–840.
23. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw* 2010; **36**: 1–48.
24. Inthout J, Ioannidis J and Borm G. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Med Res Methodol* 2014; **14**: 25. Cited By 1519.
25. Kontopantelis E and Reeves D. Performance of statistical methods for meta-analysis when true study effects are non-normally distributed: a simulation study. *Stat Methods Med Res* 2012; **21**: 409–426.
26. Partlett C and Riley RD. Random effects meta-analysis: coverage performance of 95% confidence and prediction intervals following REML estimation. *Stat Med* 2017; **36**: 301–317.
27. Doi S, Barendregt J, Khan S, et al. Advances in the meta-analysis of heterogeneous clinical trials II: the quality effects model. *Contemp Clin Trials* 2015; **45**: 123–129.
28. Lee KJ and Thompson SG. Flexible parametric models for random-effects distributions. *Stat Med* 2008; **27**: 418–434.
29. Kulinskaya E and Olkin I. An overdispersion model in meta-analysis. *Stat Model* 2014; **14**: 49–76.
30. Doi S, Barendregt J, Khan S, et al. Advances in the meta-analysis of heterogeneous clinical trials I: the inverse variance heterogeneity model. *Contemp Clin Trials* 2015; **45**: 130–138.