Meta-analysis in Stata using gllamm

Pantelis G. Bagos

There are several user-written programs for performing meta-analysis in Stata (Stata Statistical Software: College Station, TX: Stata Corp LP). These include metan, metareg, mvmeta, and glst. However, there are several cases for which these programs do not suffice. For instance, there is no software for performing univariate meta-analysis with correlated estimates, for multilevel or hierarchical meta-analysis, or for meta-analysis of longitudinal data. In this work, we show with practical applications that many disparate models, including but not limited to the ones mentioned earlier, can be fitted using gllamm. The software is very versatile and can handle a wide variety of models with applications in a wide range of disciplines. The method presented here takes advantage of these modeling capabilities and makes use of appropriate transformations, based on the Cholesky decomposition of the inverse of the covariance matrix, known as generalized least squares, in order to handle correlated data. The models described earlier can be thought of as special instances of a general linear mixed-model formulation, but to the author’s knowledge, a general exposition in order to incorporate all the available models for meta-analysis as special cases and the instructions to fit them in Stata has not been presented so far. Source code is available at http://www.compgen.org/tools/gllamm. Copyright © 2015 John Wiley & Sons, Ltd.

Keywords: meta-analysis; tutorial; Stata; gllamm

1. Introduction

The continuously increasing number of published studies in all areas of biomedical sciences made imperative the need of collecting and synthesizing the available data. The statistical analysis used to synthesize multiple studies is known as meta-analysis (Normand, 1999; Petiti, 1994; Trikalinos et al., 2008). In meta-analysis, a set of original studies is synthesized and the potential heterogeneity is explored using formal statistical methods (Glass, 1976; Greenland, 1998; Normand, 1999; Petiti, 1994). In the medical literature, meta-analysis was initially applied in the field of randomized clinical trials (Chalmers et al., 1987; Sacks et al., 1987), but nowadays, it is considered a valuable tool for the combination of observational studies (Stroup et al., 2000), as well as for genetic association studies for which specialized methodology has been developed (Salanti et al., 2007; Trikalinos et al., 2008; Bagos, 2008; Thakkinstian et al., 2005; Salanti and Higgins, 2008).

Statistical models for meta-analysis range from the simple univariate meta-analysis (DerSimonian and Laird, 1986; Normand, 1999) and meta-regression (Thompson and Higgins, 2002), to the more complex multivariate meta-analysis with its various instances (van Houwelingen et al., 2002; Berkey et al., 1998), and to the meta-analysis of correlated estimates (Gleser and Olkin, 1994); such as those appearing in the dose–response models (Greenland and Longnecker, 1992), in hierarchical models (Stevens and Taylor, 2009), and in longitudinal data (Jones et al., 2009; Ishak et al., 2007; Peters and Mengersen, 2008). All these models can be thought of as special instances of a general linear mixed-model formulation that has been proposed in the past (Platt et al., 1999; Stram, 1996), but to the author’s knowledge, a general exposition in order to incorporate all the available models for meta-analysis as special cases and the instructions to fit them in Stata has not been presented so far.

For performing meta-analysis in Stata, there are several user-written programs that facilitate the analysis providing a user-friendly interface. These include metan for univariate meta-analysis (Harris et al., 2008), metareg for meta-regression (Harbord and Higgins, 2008), mvmeta for multivariate meta-analysis (White, 2009), and glst for meta-analysis of dose–response models (Orsini et al., 2006). However, there are several cases where these programs do not suffice. For instance, there is no dedicated module for performing univariate meta-analysis with correlated estimates, or for performing a multilevel meta-analysis (i.e., when we have multiple groups nested within studies).
There is, however, a Stata program, gllamm, that can fit linear mixed models, and this can also be used for meta-analysis. gllamm (Rabe-Hesketh et al., 2002; Rabe-Hesketh et al., 2005) is a very versatile software that can handle a wide variety of models such as generalized linear mixed models, multilevel regression models, factor models, item response models, structural equation models, and latent class models with applications in a wide range of disciplines (Skrandal and Rabe-Hesketh, 2004). gllamm uses numerical integration by adaptive quadrature to integrate the latent variables and obtain the marginal log-likelihood. Afterwards, the log-likelihood is maximized by Newton–Raphson using numerical first and second derivatives. gllamm has been used in several applications in meta-analysis; most of them, however, use its remarkable features in estimating generalized linear mixed models that use directly binary or count data. Such methods, which are usually termed individual patient data (IPD) methods, are superior in many circumstances compared with the usual summary-data methods, because they avoid the normality assumptions and they do not need continuity correction in case of zero cell counts (Cooper and Patall, 2009; Stewart and Tierney, 2002). Users with no prior experience with gllamm should consult the Appendix, where several tutorials are listed as well as simple instructions for installation.

In this work, we are going to show with practical examples that many seemingly unrelated models, including but not limited to the ones mentioned in the preceding text, can be fitted using gllamm in a linear mixed-model framework. Using a regression framework, apart from the pedagogical role that it plays, allows also the full use of the relevant machinery, that is, the ability to check modeling assumptions, to detect outliers, or to perform formal model comparisons. In the following section, we will describe in detail the mathematical aspects of the method. Afterwards, we are going to present applications in univariate meta-analysis and meta-regression, multivariate meta-analysis, and meta-analysis of correlated and hierarchical data. All data and source code needed for these analyses is given in www.compgen.org/tools/gllamm and in the journal’s website.

2. Methods

In this section, we will present the mathematical framework for the method. In particular, we are going to show that gllamm can be used for performing meta-analysis, even with correlated data, and this can be achieved by an appropriate transformation of the original data. The derivation is simple, and it is provided here for completeness. However, the understanding is really not necessary for nonexperts that simply want to use gllamm for meta-analysis.

Let us consider the general case of a meta-analysis of $k$ studies, each contributing $n_i$ observations, with the total number of observations being $\sum n_i = n$. In the general case, the meta-analysis can be modeled using a linear mixed model, which in the matrix form can be represented as

$$y = X\beta + Zb + \varepsilon$$  \hspace{1cm} (1)

Here, $y$ is the $n \times 1$ response vector, $\beta$ is the $p \times 1$ vector of fixed-effects coefficients associated with $X$, which is the $n \times p$ design matrix for the fixed effects, $b$ is the $kq \times 1$ vector of random effects associated with $Z$, which is a specially designed $n \times kq$ design matrix for the random effects, and finally, $\varepsilon$ is the error. In the meta-analysis, the elements of $y$ ($y_i$) are the outcome data, extracted from the published reports, (e.g., the log(odds ratios) or any other commonly used effect size). The covariance structure can be represented by

$$b \sim N(0, G)$$  \hspace{1cm} (2)

and

$$\varepsilon \sim N(0, \Omega)$$  \hspace{1cm} (3)

The $n \times n$ matrix $\Omega$ is also made from data extracted from the published reports and is assumed to be known. We also critically assume that $\Omega$ is positive semi-definite. On the other hand, $G$ is $kq \times kq$ block-diagonal matrix, $G = diag(D, D, ..., D) = I_k \otimes D$, with $D$ being a $q \times q$ matrix that is estimated during the fitting procedure. Within a frequentist framework, the marginal model is viewed as $y \sim N(X\beta, ZGZ' + \Omega)$, so the generalized least-square (GLS) estimate of $\beta$ given known covariance parameters is given by

$$\hat{\beta} = (X'V^{-1}X)^{-1}X'V^{-1}y$$  \hspace{1cm} (4)

where $V = ZGZ' + \Omega$ and var($\beta$) = $(X'V^{-1}X)^{-1}$. The model of Eq. (1), with the definitions given in the preceding text, is quite general, and most of the commonly encountered models used for meta-analysis can be seen as special cases (we assume separate random effects for each study, but this can be relaxed later). In this formulation, we implicitly assume a univariate response (e.g., $y$ is $n \times 1$), but the model can be used for multivariate responses with some simple transformations and appropriate definition of the design matrices $X$ and $Z$. We also note that we adapted slightly the definitions given by Laird and Ware (1982), which used the model for longitudinal data, in order to accommodate all the special cases. A linear mixed-model formulation similar to Eq. (1) has been used
in the past (Platt et al., 1999; Stram, 1996). Van Houwelingen and coworkers used such a formulation for multivariate meta-analysis and meta-regression and provided illustrative code in SAS PROC MIXED (2002). Stijnen and coworkers followed a similar approach, but instead, they used the exact (binomial or Poisson) likelihood in a generalized linear mixed model, providing code in SAS PROC NLMIXED (2010). However, to the author’s knowledge, a general exposition in order to incorporate all the available models for meta-analysis as special cases and the instructions to fit them in Stata has not been presented so far.

Note, however, that the model of Eq. (1) in its general form incorporates heteroscedastic as well as correlated disturbances. gllamm includes a built-in functionality for incorporating heteroscedastic errors, but not correlated ones. Thus, as a first step, we are going to use a simple transformation in order to convert the model into one that can be fitted with gllamm. The transformation is well-known in the literature, GLS approach, which appears in all standard textbooks (Green, 2008) and is based on the Cholesky decomposition of the inverse of the covariance matrix. To illustrate our approach, we will begin by calculating $P$, the square root of $\Omega^{-1}$ using the Cholesky decomposition:

$$PP = \Omega^{-1}$$  \hspace{1cm} (5)

Clearly, $P$ satisfies

$$P\Omega P' = P\Omega(\Omega^{-1} P^{-1}) = I_n$$  \hspace{1cm} (6)

Then, we will left multiply Eq. (1), in order to obtain $Py = PX\beta + PZb + Pe$. Thus, we will now have a new model for the transformed variables:

$$\tilde{y} = \tilde{X}\beta + \tilde{Z}b + \tilde{e}$$  \hspace{1cm} (7)

We see that disturbances are standard normal, because

$$\text{var}(\tilde{e}) = \text{var}(Pe) = P \text{var}(e)P' = P\Omega P' = I_n$$  \hspace{1cm} (8)

Thus, the new covariance structure will now be $b \sim N(0, \Omega)$ and $\tilde{e} \sim N(0, I_n)$. Note that this model can easily be estimated using gllamm. The standard GLS estimate given known covariance parameters, $(\beta^*)$, will now be equal to

$$\beta^* = \left(X'X^{-1}X^{-1}V^{-1}y\right)$$  \hspace{1cm} (9)

with $V = \tilde{Z}\tilde{G}\tilde{Z}^{'1}+I_n$. Then, we will have

$$\beta^* = \left(X'X^{-1}X^{-1}V^{-1}y\right)$$

$$= \left((PX)'V^{-1}PX\right)^{-1}(PX)'V^{-1}Py$$

$$= \left(X'P^{-1}PX\right)^{-1}X'P^{-1}Py$$  \hspace{1cm} (10)

Using Eq. (5), we will have

$$V^{-1} = \left(\tilde{Z}\tilde{G}\tilde{Z}^{'1}+I_n\right)^{-1} = \left(P\tilde{Z}\tilde{G}\tilde{Z}^{'1}P + I_n\right)^{-1}$$

$$= \left(P(V - \Omega)P' + I_n\right)^{-1} = \left((P(V - \Omega)P' + I_n)^{-1}\right)$$

$$= \left(P(V - \Omega)P' + I_n\right)^{-1} = \left(P(V - \Omega)P' + I_n\right)^{-1}$$

$$= \left(P(V - \Omega)P' + I_n\right)^{-1} = \left(P(V - \Omega)P' + I_n\right)^{-1}$$  \hspace{1cm} (11)

From Eqs. (10) and (11), we obtain

$$\beta^* = \left(X'P^{-1}PX\right)^{-1}X'P^{-1}Py$$

$$= \left(X'X^{-1}X^{-1}V^{-1}y\right)$$  \hspace{1cm} (12)

Moreover, from Eqs. (10) and (12), it is also obvious that $\text{var}(\beta^*) = \left(X'X^{-1}X^{-1}\right) = \left(X'X^{-1}X^{-1}\right) = \text{var}(\beta)$. In other words, the fixed-effects estimates and their variances obtained by fitting the model of Eq. (7) are identical to the ones of the model of Eq. (1), as one would expect. The random-effects parameters $(D \text{ or } G)$, which are usually estimated using maximum likelihood (ML), are also identical following the general likelihood property stating that likelihood estimation is invariant under a one-to-one transformation. Thus, we can use a program
capable of fitting linear mixed models with uncorrelated errors, such as gllamm, for fitting the model of Eq. (1); we only need to transform the variables appropriately.

Note that, when $\Omega$ is diagonal, $\Omega = \text{diag}(s_1^2, s_2^2, \ldots, s_k^2)$; that is, when the errors are independent, the situation is simplified because the model reduces to the weighted least-square approach, with weights given by the inverse of the square root of the variance. Then,

$$\Omega^{-1} = w^tw$$ with $w = \begin{bmatrix} 1/s_1 & 0 & \ldots & 0 \\ 0 & 1/s_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1/s_n \end{bmatrix}$ \hspace{1cm} (13)

This transformation, known as weighted least squares, is a special case of the GLS approach outlined previously. If, additionally, the model contains a random intercept, no covariates, and only one observation per study, then we are talking about the special case of univariate random-effects meta-analysis. In this case, $n_i = 1$ for $i = 1, 2, \ldots, k$ (and thus $n = k$), $\beta$ and $D$ are scalars, $y$ and $X$ are $k \times 1$ vectors ($X$ is a vector of 1’s, denoted by $1_k$), and $\Omega$ is $k \times k$. From these, we have that $Z = 1_k$ and $G = 1_k \otimes D = r^2 1_k$. With these remarks, Eq. (4) reduces to

$$\beta = \frac{1V^{-1}y}{1V^{-1}1}$$

with $V$ given by

$$V = ZG' + \Omega = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix} \begin{bmatrix} r^2 & 0 & \ldots & 0 \\ 0 & r^2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & r^2 \end{bmatrix} \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix} + \begin{bmatrix} s_1^2 & 0 & \ldots & 0 \\ 0 & s_2^2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & s_k^2 \end{bmatrix}$$

Moreover, we will have

$$1V^{-1}y = \begin{bmatrix} 1/(s_1^2 + r^2) & 0 & \ldots & 0 \\ 0 & 1/(s_2^2 + r^2) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1/(s_k^2 + r^2) \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix} = \sum_{i=1}^k (s_i^2 + r^2)^{-1} y_i$$

$$1V^{-1}1 = \begin{bmatrix} 1/(s_1^2 + r^2) & 0 & \ldots & 0 \\ 0 & 1/(s_2^2 + r^2) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1/(s_k^2 + r^2) \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \sum_{i=1}^k (s_i^2 + r^2)^{-1}$$

In other words, Eq. (4) reduces to the well-known inverse variance-weighted estimate:

$$\beta = \frac{\sum_{i=1}^k (s_i^2 + r^2)^{-1} y_i}{\sum_{i=1}^k (s_i^2 + r^2)^{-1}}$$

with variance given by

$$\text{var}(\beta) = (1V^{-1}1)^{-1} = \frac{1}{\sum_{i=1}^k (s_i^2 + r^2)^{-1}}$$
In short, we have shown that performing the transformation of Eq. (5) and fitting the model of Eq. (7) provide the means to perform a meta-analysis with correlated and heteroscedastic errors in a linear mixed-model framework. The estimates and their standard errors are identical, and in the special case of the univariate meta-analysis, the estimate reduces to the well-known inverse variance estimate. What it remains now is to show how the various models mentioned in the introduction section can be formulated in a linear mixed-model framework and the details of fitting such models in gllamm; this is explained in the next section. Users who are not familiar with gllamm should consult the Appendix, where several tutorials are listed and simple instructions for installation are given.

3. Application of the method

3.1. Univariate meta-analysis and meta-regression

Let us consider in more detail the simple case of a univariate meta-analysis, in which independent estimates from \( k \) studies are synthesized in order to obtain a pooled estimate (Normand, 1999). In such case, the random-effects meta-analysis model is usually described by a linear model:

\[
y_i = \beta_0 + b_i + \varepsilon_i, \quad b_i \sim N(0, \tau^2), \quad \varepsilon_i \sim N(0, s_i^2)\tag{14}
\]

where \( \beta_0 \) is the grand mean that we wish to estimate, \( b_i \) the random deviation from the average intercept associated with study \( i \), and \( \varepsilon_i \) the random error. Note that an alternative notation for error term of the model of Eq. (14) is \( \varepsilon_i \sim N(0, \Omega) \), with

\[
\Omega = \text{diag}(s_1^2, s_2^2, \ldots, s_k^2) = \begin{bmatrix}
s_1^2 & 0 & \cdots & 0 \\
0 & s_2^2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & s_k^2
\end{bmatrix}\tag{15}
\]

We usually assume that the random intercept is normally distributed with zero mean and variance \( \tau^2 \) (i.e., the heterogeneity parameter or the between-studies variance). We also critically assume, as we noted earlier, that the variances \( s_i^2 \) are known observable quantities. Sometimes, the model of Eq. (14) can be seen in the more convenient form of the marginal model:

\[
y_i \sim N(\beta_0, s_i^2 + \tau^2)\tag{16}
\]

In the simplest form, as we described previously, the estimate is simply a weighted average of the individual estimates:

\[
\hat{\beta} = \frac{\sum_{i=1}^k w_i y_i}{\sum_{i=1}^k w_i}, \quad w_i = \frac{1}{s_i^2 + \tau^2}\tag{17}
\]

Traditionally, the random-effects parameter, \( \tau^2 \), is estimated using the method of moments of DerSimonian and Laird (1986), even though methods based on ML or restricted ML (REML) are also available (Thompson and Sharp, 1999). In case there is no heterogeneity (\( \tau^2 = 0 \)), the model reduces to the well-known inverse variance method of Woolf (1955). In the previous section, we have shown that this estimate is identical to the one produced by the linear mixed model. We only need to have in mind that gllamm uses exclusively ML as the method for estimating \( \tau^2 \); of course, we expect that in case of fixed-effects model, the results will be identical. In Stata, traditionally the `metan` command is used to implement this method with a variety of options for calculating the weights, displaying graphically the results, as well as allowing for different effect sizes (Harris et al., 2008). However, it only allows Method of moments (MM) estimates for \( \tau^2 \).

In order to begin the exposition of the method, it is necessary to remind that exactly the same fixed-effects estimate can be obtained, using an ordinary regression command, with the exception that the mean square error (MSE) has to be forced to be unity (Thompson and Sharp, 1999). The correct standard errors of the regression coefficients are thus obtained by dividing those obtained by the regression program by the square root of the reported MSE.

We illustrate the method in the do-file `univariate.do` that contains the data and the commands for replicating the meta-analysis of clinical trials on the efficacy of Bacillus Calmette-Guérin (BCG) vaccine in the prevention of tuberculosis (Colditz et al., 1994). These data were also used by van Houwelingen (van Houwelingen et al., 2002). The meta-analysis concerns 13 trials on the efficacy of BCG vaccine against tuberculosis, and in each trial, a vaccinated group is compared with a non-vaccinated control group. Several covariates are available that might explain the heterogeneity among studies, and following van Houwelingen, we will present the results using the geographic latitude of the place where the study was performed.
The command to perform the univariate fixed-effects meta-analysis (\texttt{metan}) using inverse variance weights is as follows:

\begin{verbatim}
.metan logor se, fixed
\end{verbatim}

The exact same result can be reproduced by manually weighting the dependent variable by dividing by $s_i$, $y_i = y_i/s_i$, creating a dummy variable $\tilde{x}_i$ by dividing 1 by $s_i$, and then performing an unweighted linear regression through the origin (i.e., with no constant term) forcing, however, the MSE to be unity. In other words, we obtain the estimates by fitting the model:

$$\tilde{y}_i = \beta_0 \tilde{x}_i + \tilde{e}_i, \tilde{e}_i \sim N(0, 1)$$

The model can be fitted using \texttt{glm} as follows:

\begin{verbatim}
.gen wlogor=logor/se 
.gen w1=1/se 
.glm wlogor w1, nocons scale(1)
\end{verbatim}

For the random-effects model, we can use the command \texttt{metan}, which allows the calculation of DerSimonian and Laird (DSL) between studies heterogeneity, or alternatively, we may use the \texttt{metareg} command (Harbord and Higgins, 2008) or the \texttt{mvmeta} command (White, 2009), which both allow for ML and REML estimates. \texttt{gllamm} has also a built-in functionality that allows fitting the model of Eq. (14) and performs a weighted regression forcing the level-1 variances to an arbitrary value. This is achieved by using the logarithm of the standard errors and setting the appropriate constraints:

\begin{verbatim}
.gen logs=log(sqrt(var)) 
.eq wgt: logs 
.constraint define 1 [lns1]logs=1 
.gllamm logor, i(trial) nrf(1) eqs(logs) s(wgt) constraint(1) adapt nip(24)
\end{verbatim}

Note that \texttt{gllamm} allows the level-1 variance to depend on covariates through a linear predictor with no intercept. By fixing the coefficient of log($s_i$) to be 1, we ensure that the variance is fixed to the given values, because \texttt{gllamm} internally exponentiates the covariates for numerical reasons. The \texttt{nip} option controls the number of integration points used by adaptive quadrature; larger numbers lead to more accurate results but slow down the computation. As more random effects are added to the model, more integration points are required, but for the majority of models with one or two random effects, usually 12 integration points are adequate. Moreover, by repeating the last \texttt{gllamm} command appending in the end the \texttt{init} option (initialization), we can recover the same fixed-effects model fitted by \texttt{glm} (the results are identical).

\begin{verbatim}
.gllamm logor, i(trial) nrf(1) eqs(logs) s(wgt) constraint(1) adapt nip(24)init
\end{verbatim}

In order to fit the same model using the manual weighting technique, we need to fit a random coefficient model of the form

$$\tilde{y}_i = \beta_0 \tilde{x}_i + b \tilde{x}_i + \tilde{e}_i, \tilde{e}_i \sim N(0, \tau^2), \tilde{e}_i \sim N(0, 1)$$

This model can be fitted with \texttt{gllamm} using

\begin{verbatim}
.gen constant=1 
.eq cons:constant 
.eq slope:w1 
.constraint define 2 [lns1]cons=0 
.gllamm wlogor w1, nocons i(trial) nrf(1) constraint(2) eq(slope) s(cons) adapt nip(24)
\end{verbatim}

Note that with this syntax, the level-1 variance needs to be set to one, and hence, the coefficient for the logarithm of the standard deviation must be zero. Moreover, we need a specific equation for the random coefficient (i.e., the slope), and we must ensure that the model contains no intercept, neither fixed nor random. The random-effects estimates obtained by the two alternative \texttt{gllamm} syntaxes in the preceding text are nearly identical (with agreement up to the fifth decimal place) but differ slightly compared with the results obtained by \texttt{metan}. This is because \texttt{metan} uses the DSL method of moments, whereas \texttt{gllamm} uses ML. On the contrary, these results should be compared with the ones obtained by \texttt{metareg} and \texttt{mvmeta}, as well as by a hand-written Stata program that uses \texttt{ml} (we provide such a program in the supplement). In Table 1, we list the results...
Table 1. Results obtained for the univariate meta-analysis on the BCG data using the various methods described in text.

<table>
<thead>
<tr>
<th>Method</th>
<th>( \beta_0 ) (SE)</th>
<th>( \tau^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>gllamm (syntax 1)</td>
<td>-0.741979 (0.178647)</td>
<td>0.302510</td>
</tr>
<tr>
<td>gllamm (syntax 2)</td>
<td>-0.741979 (0.178665)</td>
<td>0.302587</td>
</tr>
<tr>
<td>metareg (ML)</td>
<td>-0.741967 (0.177953)</td>
<td>0.302393</td>
</tr>
<tr>
<td>mvmeta (ML)</td>
<td>-0.741967 (0.178617)</td>
<td>0.302456</td>
</tr>
<tr>
<td>Stata ml</td>
<td>-0.747392 (0.192262)</td>
<td>0.366343</td>
</tr>
</tbody>
</table>

ML, maximum likelihood; SE, standard error.

Table 2. Results obtained on the univariate meta-regression of latitude on the log-odds ratio for the BCG data.

<table>
<thead>
<tr>
<th>Method</th>
<th>( \beta_0 ) (SE)</th>
<th>( \beta_1 ) (SE)</th>
<th>( \tau^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>gllamm (method 1)</td>
<td>0.371049 (0.131540)</td>
<td>-0.032722 (0.003580)</td>
<td>0.003999</td>
</tr>
<tr>
<td>gllamm (method 2)</td>
<td>0.371141 (0.131512)</td>
<td>-0.032725 (0.003578)</td>
<td>0.003989</td>
</tr>
<tr>
<td>metareg (ML)</td>
<td>0.372509 (0.104389)</td>
<td>-0.032745 (0.003333)</td>
<td>0.003581</td>
</tr>
<tr>
<td>mvmeta (ML)</td>
<td>0.371070 (0.131526)</td>
<td>-0.032722 (0.003579)</td>
<td>0.003998</td>
</tr>
<tr>
<td>Stata ml</td>
<td>0.371070 (0.131526)</td>
<td>-0.032722 (0.003579)</td>
<td>0.003998</td>
</tr>
</tbody>
</table>

ML, maximum likelihood; SE, standard error.
is metareg, as one would expect based on the fact that the approximate algorithm that it uses has been found to be not so accurate in several applications.

3.2. Multivariate meta-analysis

In the simple case of univariate meta-analysis presented in the previous section, the parameter of interest was one-dimensional. In many applications, however, we need to consider simultaneously in a meta-analysis two or more estimates arising from the same study. These situations may occur in different settings, when the parameter of interest is truly multivariate, such as with multiple treatments (Higgins and Whitehead, 1996) or multiple outcomes (Gleser and Olkin, 1994), or in genetic association studies (Bagos, 2008; Minelli et al., 2005). However, such a situation may emerge in other instances such as in meta-analysis of diagnostic tests (Arends et al., 2008a; Harbord et al., 2007; Reitsma et al., 2005) and in meta-analysis adjusting for the baseline risk (Thompson et al., 1997).

In any case, in multivariate meta-analysis, we are interested in performing a joint modeling of the different quantities of interest (van Houwelingen et al., 2002; Mavridis and Salanti, 2012, Jackson et al., 2011). The multivariate meta-analysis enables the researcher to perform global tests avoiding multiple comparisons and consequently, the inflation of type I error rate. Additionally, we can estimate the covariance matrix for all the parameters of interest. This could be of interest by itself, for instance, when it comes to calculating the correlation between a surrogate and a true outcome, but it can also be of importance when we want to perform a comparison of the two estimates and construct a confidence interval for the difference or some other functions of the estimates (Altman and Bland, 2003). Finally, a major advantage of multivariate meta-analysis is that it can accommodate, under the missing-at-random assumption, studies reporting only one of the parameters of interest, resulting in borrowing strength from external studies (Higgins and Whitehead, 1996).

In any case, if the estimates are uncorrelated within studies, the situation is straightforward. However, it becomes complicated when the estimates are stochastically dependent, in which case, the within-study correlation and the associated covariance need to be calculated. This, however, in most situations requires access to the individual data that may not be available to a researcher performing the analysis, but in several special cases, this covariance can be calculated analytically (Bagos, 2012).

In the following, we are going to describe the bivariate meta-analysis, but the generalization to more than two outcomes is trivial. Following the general framework for multivariate meta-analysis, we denote by \( y \), the vector containing the two estimates and by \( \beta \) the vector of the overall means:

\[
y_i = \begin{bmatrix} y_{1i} \\ y_{2i} \end{bmatrix} \quad \text{and} \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}
\]

We further assume that \( y_i \) is distributed following a multivariate normal distribution around the true means \( \beta \), according to the marginal model:

\[
y_i \sim \text{MVN}(\beta, D + \Sigma_i)
\]

By \( \Sigma_i \) we denote the within-studies covariance matrix:

\[
\Sigma_i = \begin{bmatrix}
s_{11}^2 & \rho_{W_1S_{11}} s_{21}
\rho_{W_1S_{21}} & s_{21}^2
\end{bmatrix}
\]

The diagonal elements of \( \Sigma_i \) are the study-specific estimates of the variance that are considered known, whereas the off-diagonal elements correspond to the pairwise within-studies covariances, for instance, \( \rho_{W_1S_{11}} = \text{cov}(y_{2i}, y_{1i}) \). By \( D \) we denote the between-studies covariance matrix, which is to be estimated during the fitting process:

\[
D = \begin{bmatrix}
t_1^2 & \rho_{B_1 \tau_2}
\rho_{B_1 \tau_2} & t_2^2
\end{bmatrix}
\]

Thus, the final marginal model on which we base the inference is:

\[
\begin{bmatrix} y_{1i} \\ y_{2i} \end{bmatrix} \sim \text{MVN} \left( \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}, \begin{bmatrix}
s_{11}^2 + t_1^2 & \rho_{W_1S_{11}} s_{21} + \rho_{B_1 \tau_2} t_2 \\
\rho_{W_1S_{21}} s_{21} + \rho_{B_1 \tau_2} t_2 & s_{21}^2 + t_2^2
\end{bmatrix} \right)
\]

In order to fit such models using gllamm, we will need to properly rearrange the data to meet the requirements of Eq. (1). That is, we need to stack the multiple outcomes in a single column and subsequently, create indicator variables for each outcome, which will give \( X \).
The model can then be written in the form of a random coefficient linear model with no intercept:

\[ y_{ij} = \beta_1 x_{1i} + b_{1i} x_{1i} + \beta_2 x_{2i} + b_{2i} x_{2i} + \epsilon_{ij}, \quad b_i \sim N(0, D), \epsilon_{ij} \sim N(0, \Omega) \]  

with between-studies covariance matrix:

\[
D = \begin{pmatrix} \tau_1^2 & 2 \rho_{B1\tau_2} \tau_2 & \tau_2^2 & \\ 2 \rho_{B1\tau_2} \tau_2 & \tau_2^2 & \tau_2^2 & \\ \tau_2^2 & \tau_2^2 & \tau_2^2 & \tau_2^2 \end{pmatrix}, \quad \text{and } \Omega = diag(\Sigma_1, \Sigma_2, \ldots, \Sigma_k) = \begin{pmatrix} \Sigma_1 & 0 & \cdots & 0 \\ 0 & \Sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Sigma_k \end{pmatrix} 
\]

The full between-studies covariance matrix \( D \) will be given by

\[
G = diag(D, D, \ldots, D) = I_k \otimes D = \begin{pmatrix} D & 0 & \cdots & 0 \\ 0 & D & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D \end{pmatrix} 
\]

Because we have a random coefficient model, \( X \) will help to determine \( Z \). We now have two rows for each \( i = 1, 2, \ldots, k \), and two random effects, so \( x_i = I_2 \), and thus, we will have

\[
Z = \begin{pmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_k \end{pmatrix} = \begin{pmatrix} I_2 & 0 & \cdots & 0 \\ 0 & I_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I_2 \end{pmatrix} 
\]

and finally

\[
V = ZGZ' + \Omega = \begin{pmatrix} I_2 & 0 & \cdots & 0 \\ 0 & I_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I_2 \end{pmatrix} = \begin{pmatrix} D + \Sigma_1 & 0 & \cdots & 0 \\ 0 & D + \Sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D + \Sigma_k \end{pmatrix} = diag(D + \Sigma_1, D + \Sigma_2, \ldots, D + \Sigma_k) 
\]
That is, the model of Eq. (28) has the same marginal interpretation as the model of Eq. (26). Using this formulation, \( y \) is now \( 2k \times 1 \), and so \( \Omega, G, Z, \) and \( V \) will be \( 2k \times 2k \).

### 3.2.1. Multivariate meta-analysis for baseline risk (no within-studies correlation)

Following van Houwelingen et al. (2002), we will begin the introduction to the bivariate approach with special reference to the situation where one is interested in “control rate regression”, that is, the approach that tries to relate the treatment effect size to the risk of events in the control group. However, this approach applies generally, and generalization to the multivariate case with more than two outcomes is trivial. Many studies show considerable variation in what is called the baseline risk, which indicates the risk for patients in the control group. One might wonder if there is a relation between treatment effect and baseline risk because considering only the differences between the study arms may hide much information (Thompson et al., 1997; van Houwelingen and Senn, 1999). There are of course other ways of performing a similar analysis, for instance, by modeling simultaneously the log (odds ratio) and the baseline log(odds). However, these will be correlated, and thus, the method described in the next section will be needed. Therefore, it may be wise to consider the pair of outcomes of the two treatments. In this case, the within-studies covariances are zero, and thus, the within-studies covariance matrix simplifies to

\[
\Omega = \text{diag}(s_{11}^2, s_{21}^2, \ldots, s_{1k}^2, s_{2k}^2)
\]

We, once again, use the data on BCG vaccination (bcg.dta). The commands that transform the data and fit the model using gllamm with syntax 1 are given in the following (they are also given in the bivariate.do file):

```stata
.reshape long b var, i(trial) j(trt) string
.quietly tabulate trt, generate(grp)
.eq grp1: grp1
.eq grp2: grp2
.gen sd=log(sqrt(var ))
.eq wgt: sd
.constraint define 1 [lns1]sd=1
gllamm b grp1 grp2, nocons i(trial) nrf(2) eqs(grp1 grp2) s(wgt) constraint(1) nip(16) adapt
```

We need to emphasize that the model of Eq. (28) and all the multivariate models that follow need the data in the so-called “long format” (in Stata jargon). Thus, the transformation of Eq. (27) needs to be carried out prior to issuing the gllamm statements, and this is achieved using the reshape command. In order to use the second method (syntax 2), we need to further transform the data:

```stata
.gen cons=1
.gen wb=b/sqrt(var)
.gen wgrp1=grp1/sqrt(var)
.gen wgrp2=grp2/sqrt(var)
.eq wgrp1: wgrp1
.eq wgrp2: wgrp2
.eq cons: cons
.constraint define 2 [lns1]cons=0
gllamm wb wgrp1 wgrp2, nocons i(trial) nrf(2) eqs(wgrp1 wgrp2) s(wgt) constraint(2) s(cons) adapt nip(16)
```

The results are listed in Table 3, where we can see that all methods provide identical estimates. Their interpretation follows the excellent tutorial of van Houwelingen et al. (2002) and is as follows. The fixed-parameter estimates (−4.834, −4.096) represent the estimated mean log-odds in the vaccinated and non-vaccinated group. The between-study estimated variance of the log-odds is 1.431 in the vaccinated groups and 2.407 in the non-vaccinated group.
not-vaccinated groups. The between-study covariance is estimated to be 1.757. Thus, the estimated correlation between the true vaccinated and true control log-odds is 0.947. With the bivariate approach, we can easily recover using the delta method the results of the univariate analysis.

\[ \text{lincom} \ _b[\text{grp2}]-_b[\text{grp1}] \]

The estimated mean vaccination effect, measured as the log-odds ratio, is equal to \((-4.834 - (-4.096)) = -0.738\). The standard error of the mean vaccination effect is equal to \((0.115 + 0.189 \cdot 2 \cdot 0.136) = 0.181\), slightly larger compared with the result of the univariate model. This corresponds to an estimated odds ratio of \(\exp(-0.738) = 0.478\) (95\% confidence interval [CI]: 0.335, 0.682), again strongly suggesting an average beneficial vaccination effect.

However, the bivariate model can be more useful because it allows the calculation of the slope of the regression line to predict the log-odds in the vaccinated group from the log-odds in the not-vaccinated group, a calculation that yields \(1.757/2.407 = 0.730\).

\[ \text{nlcom} \ (\text{tri1}_2\text{.1}\_\text{cons}*\text{tri1}_2\text{.1}\_\text{cons} + \text{tri1}_2\text{.2}\_\text{cons}*\text{tri1}_2\text{.2}\_\text{cons} -2*\text{tri1}_2\text{.1}\_\text{cons}*\text{tri1}_1\_\text{grp1})/(	ext{tri1}_1\_\text{grp1}^2) \]

Similarly, we can calculate the variance of the treatment effect, measured as the log-odds ratio, a calculation that yields \(1.431 + 2.407/2 \cdot 1.757 = 0.324\), which is only slightly different from what we found earlier in the univariate random-effects analysis.

\[ \text{nlcom} \ [\text{tri1}_1\_\text{grp1}]*[\text{tri1}_1\_\text{grp1}]+[\text{tri1}_2\_\text{cons}]*[\text{tri1}_2\_\text{cons}]*[\text{tri1}_2\_\text{cons}]*[\text{tri1}_1\_\text{grp1}]+[\text{tri1}_2\_\text{grp2}]*[\text{tri1}_2\_\text{grp2}]-2*[\text{tri1}_2\_\text{cons}]*[\text{tri1}_1\_\text{grp1}] \]

Finally, we can obtain the conditional variance of the true log-odds, and therefore also of the log-odds ratio, in the vaccinated group given the true log-odds in the not-vaccinated group \((1.431 - 1.757)/2.407 = 0.149\), which is interpreted as the variance between treatment effects among studies with the same baseline risk. So, baseline risk, measured as the true log-odds in the not-vaccinated group, explains \(0.324 - 0.149)/0.324 = 54.2\%\) of the heterogeneity in vaccination effect between the trials.

\[ \text{nlcom} \ [\text{tri1}_2\_\text{cons}]*[\text{tri1}_2\_\text{cons}]+[\text{tri1}_2\_\text{grp2}]*[\text{tri1}_2\_\text{grp2}]-2*[\text{tri1}_2\_\text{cons}]*[\text{tri1}_1\_\text{grp1}] \]

We need to emphasize that the same model can be used for performing meta-analysis of diagnostic studies, in which case, we can recover the Summary Receiver Operator Characteristic (SROC) method (Chu and Cole, 2006; Arends et al., 2008a; Harbord et al., 2007; Reitsma et al., 2005). As a matter of fact, the two available Stata commands for fitting such models \text{metandi} (Harbord and Whiting, 2009) and \text{midas} (Dwamena, 2007) use internally \text{gllamm} as one of the methods of estimation. The do-file \text{roc.do} contains the necessary commands for performing a meta-analysis of diagnostic tests using \text{gllamm}. The relevant data are stored in the \text{roc.dta} file.

### Table 3. Results obtained for the bivariate meta-analysis on the BCG data using the various methods described in text.

<table>
<thead>
<tr>
<th>Method</th>
<th>(\beta_1) (SE)</th>
<th>(\beta_2) (SE)</th>
<th>(\tau_1^2)</th>
<th>(\tau_2^2)</th>
<th>(\rho)</th>
</tr>
</thead>
<tbody>
<tr>
<td>\text{gllamm} (method 1)</td>
<td>-4.095975 (0.434745)</td>
<td>-4.833744 (0.340198)</td>
<td>2.407333</td>
<td>1.431371</td>
<td>0.946691</td>
</tr>
<tr>
<td>\text{gllamm} (method 2)</td>
<td>-4.095975 (0.434745)</td>
<td>-4.833744 (0.340198)</td>
<td>2.407344</td>
<td>1.431371</td>
<td>0.946691</td>
</tr>
<tr>
<td>\text{mvmeta} (ML)</td>
<td>-4.095975 (0.434804)</td>
<td>-4.833744 (0.339965)</td>
<td>2.407333</td>
<td>1.431371</td>
<td>0.946691</td>
</tr>
<tr>
<td>\text{Stata ml}</td>
<td>-4.095975 (0.434745)</td>
<td>-4.833744 (0.340198)</td>
<td>2.407333</td>
<td>1.431371</td>
<td>0.946691</td>
</tr>
</tbody>
</table>

ML, maximum likelihood; SE, standard error.
3.2.2. Multiple outcomes and multiple treatments (with within-studies correlation). In some other applications, such as with multiple outcomes (Berkey et al., 1998; Berkey et al., 1995) or multiple treatments (Higgins and Whitehead, 1996), the response is truly multivariate, and the estimates are stochastically dependent. Thus, the model of Eq. (28) is combined with a within-studies covariance matrix of the form

\[
\Omega = \begin{bmatrix}
    s_{11} & \rho_{W1} & s_{12} & \cdots & s_{1k} \\
    \rho_{W1} & s_{21} & \rho_{W2} & \cdots & \rho_{W1} \rho_{W2} \\
    s_{12} & \rho_{W2} & s_{22} & \cdots & \rho_{W2} s_{2k} \\
    \cdots & \cdots & \cdots & \cdots & \cdots \\
    s_{1k} & \rho_{W1} \rho_{W2} & \rho_{W2} s_{2k} & \cdots & s_{2k}
\end{bmatrix}
\]

In this case, the within-studies correlation may arise from different reasons. For instance, when we compare multiple treatments, the correlation arises because of the existence of a common comparison group (Higgins and Whitehead, 1996). In multiple outcomes, usually the correlation arises because we have measurements on the same subjects (Berkey et al., 1998; Berkey et al., 1995). Other more complicated cases may arise in bivariate meta-analysis of genetic association studies (Bagos, 2008), when we have overlapping sets of subjects (Lin and Sullivan, 2009; Bagos et al., 2011), mutually exclusive outcomes (Trikalinos and Olkin, 2008), or multipoint meta-analysis in genetics (Bagos and Liakopoulos, 2010). In the traditional approach, the within-studies correlation needs to be provided by the individual studies (Berkey et al., 1998; Berkey et al., 1995), but recent studies have shown that for a wide range of possible designs, these correlations can be calculated analytically or imputed using the correlation between the outcomes of interest (Bagos, 2012; Wei and Higgins, 2012).

In the first example, we use the data analyzed by Berkey et al. (1998). This meta-analysis concerns five randomized controlled trials, where a surgical procedure for the treatment of periodontitis is compared with a nonsurgical procedure. Two outcomes are assessed: change in probing depth and change in attachment level. The data were originally published by Antczak-Bouckoms et al. (1993). The data are given in the file `periodontitis.dta`. In this example, because we have nonzero within-studies correlation, only syntax 2 is available (see also the do-file `periodontitis.do`). The necessary commands are given in the following:

```
. reshape long b v, i(trial) j(trt) string
. drop if b==.
. mkm v,mat(Vd)
. mat V=diag(Vd)
. tab trt,gen(x)
. local obs=_N
. forvalues x=1(2) `obs’ {
    . mat V[`x’+1,’x’]=V12[‘x’]
    . mat V[‘x’, `x’+1]=V12[‘x’]
}
. mat W=cholesky(invsym(V))
. mat P=W'
. mkm b,mat(Y)
. mat wb=P*Y
. svmat wb
. mkm x1 x2, mat(X)
. mat PX=P*X
. svmat PX
. gen cons=1
. eq cons:cons
. eq wgrp1: PX1
. eq wgrp2: PX2
. constraint define 2 [lns1]cons=0
. gllamm wb1 PX1 PX2, nocons(i(trial) nrf(2) eqs(wgrp1 wgrp2) constraint(2) s(cons) adapt)
```

As we already mentioned, `gllamm`, contrary to `mvmeta`, needs the data in “long format”, so in this case, apart from reshaping the data, we also need to generate the between-studies covariance matrix of Eq. (29) and then to perform the Cholesky decomposition of Eq. (5) and the transformation of Eq. (7). More details can be found in the
respective do-files. The results of this analysis are listed in Table 4. We also present the results obtained using mvmeta and Stata ml. It is noteworthy, that all methods show a remarkable agreement and agree also with the results obtained by Berkey et al., 1998.

The case of multiple treatments is exemplified using data from a meta-analysis of 26 trials performed to assess the effectiveness of beta-blockers and endoscopic sclerotherapy in the prevention of first bleeding in patients with cirrhosis (Pagliaro et al., 1992). Of the 26 trials, 7 compared beta-blockers against control, 17 compared sclerotherapy against control, and 2 trials compared both beta-blockers and sclerotherapy against control treatment (multi-arm trials). This dataset was used by Higgins and Whitehead (1996) in order to demonstrate the approach that leads to what is known as borrowing strength from external studies, in which the multivariate approach allows for greater precision in the estimation of the comparison of treatments (even though only a few studies compare the two treatments directly). This is a special case of a much broader area, in which we need to compare different treatments for the same disease and we are usually interested in estimating the effectiveness of each one compared with the others. The problem arises because many studies report results of only a subset of the treatments (i.e., A vs. B, B vs. C, and A vs. placebo), and thus, we need methods for integrating evidence from direct as well as from indirect comparisons. This is the so-called problem of mixed treatment comparisons leading to network meta-analysis (Gleny et al., 2005; Salanti et al., 2008). The first attempts to combine evidence for multiple treatments and perform indirect comparisons in meta-analysis date back to the 90s (Bucher et al., 1997). However, advanced multivariate techniques based on the concept of “borrowing strength from external studies” (Higgins and Whitehead, 1996) have been developed later, leading to the concept of network meta-analysis (Lumley, 2002), and various approaches have been proposed since then for mixed treatment comparison meta-analysis (Salanti et al., 2008). The data can be found in the sclerotherapy.dta file, whereas the Stata commands in the sclerotherapy.do file.

```
.reshape long b v, i(id) j(trt) string
.drop if b==.
.mkmat v, mat(Vd)
.mat V=diag(Vd)
.replace V12=0 if V11==. |V22==.
.tab trt,gen(x)
.local obs=_N
.forvalues x=1(2)`obs'{
  .mat V[`x'+1, `x'']=V12[`x'
  .mat V[`x', `x'+1]=V12[`x'
}
.mat W=cholesky(invsym(V))
.mat P=W'
.mkmat b,mat(Y)
.svmat wb
.mkmat x1 x2, mat(X)
.mat PX=P*X
.svmat PX
.gen cons =1
.eq cons:cons
.eq wgrp1: PX1
.eq wgrp2: PX2
.constraint define 2 [lns1]cons=0
.gllamm wb1 PX1 PX2, nocons i(id) nrf(2) eqs(wgrp1 wgrp2) constraint(2) s(cons)
    adapt nip(16)
```

Table 4. Results obtained for the bivariate meta-analysis on the periodontitis data using the various methods described in text.

<table>
<thead>
<tr>
<th>Method</th>
<th>( \beta_1 ) (SE)</th>
<th>( \beta_2 ) (SE)</th>
<th>( \tau_1^2 )</th>
<th>( \tau_2^2 )</th>
<th>( \rho )</th>
</tr>
</thead>
<tbody>
<tr>
<td>gllamm (method 1)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>gllamm (method 2)</td>
<td>0.344839 (0.053631)</td>
<td>—0.337938 (0.081248)</td>
<td>0.007002</td>
<td>0.026144</td>
<td>0.699230</td>
</tr>
<tr>
<td>mvmeta (ML)</td>
<td>0.344839 (0.053617)</td>
<td>—0.337938 (0.081242)</td>
<td>0.007002</td>
<td>0.026144</td>
<td>0.699230</td>
</tr>
<tr>
<td>Stata ml</td>
<td>0.344839 (0.053631)</td>
<td>—0.337938 (0.081248)</td>
<td>0.007002</td>
<td>0.026144</td>
<td>0.699230</td>
</tr>
</tbody>
</table>

Note that for this dataset, syntax 1 is not applicable.
ML, maximum likelihood; SE, standard error.
The results of this analysis are shown in Table 5. Once again, all the available methods (gllamm, mvmeta, and Stata ml) show a remarkable agreement. A sometimes useful simplification can be made in Eq. (29) if we assume that the between-studies variances are equal (Higgins and Whitehead, 1996). In such case, by letting \( \tau_1 = \tau_2 \), \( D \) reduces to

\[
D = \begin{bmatrix}
\tau^2 & \tau^2/2 \\
\tau^2/2 & \tau^2
\end{bmatrix}
\]  

(31)

Similar assumptions hold for the general case with more than two treatments. This is particularly useful in comparison of several treatments and in cases where the direct comparisons are scarce, because with this parameterization, the total number of estimated parameters is reduced. Moreover and most importantly, this parameterization is necessary in order to preserve the transitivity and consistency relations necessary for mixed treatment comparisons. The assumption of equal between-studies variances can sometimes be unrealistic, and some alternatives have been presented (Lu and Ades, 2009), but they cannot be easily applied in standard software. In order to apply this parameterization in gllamm, we note that gllamm uses the Cholesky decomposition for the between-studies covariance matrix. This consists of decomposing the covariance matrix into the product of a lower triangular matrix and its conjugate transpose, that is, \( D = LL' \). In the case of bivariate meta-analysis (with two treatments), the decomposition yields

\[
L = \begin{bmatrix}
\tau & 0 \\
\tau/2 & \tau\sqrt{3}/2
\end{bmatrix}
\]  

(32)

Similarly, for the trivariate case, we will have

\[
L = \begin{bmatrix}
\tau & 0 & 0 \\
\tau/2 & \tau\sqrt{3}/2 & 0 \\
\tau/2 & \tau\sqrt{3}/2 & \tau\sqrt{2}/3
\end{bmatrix}
\]  

(33)

and for the tetravariate,

\[
L = \begin{bmatrix}
\tau & 0 & 0 & 0 \\
\tau/2 & \tau\sqrt{3}/2 & 0 & 0 \\
\tau/2 & \tau\sqrt{3}/2 & \tau\sqrt{2}/3 & 0 \\
\tau/2 & \tau\sqrt{3}/2 & \tau\sqrt{2}/3 & \tau\sqrt{5}/2
\end{bmatrix}
\]  

(34)

The constraints of Eq. (32) can be directly applied into gllamm as follows:

```
.constraint define 3 [id1_2]PX2=-.8660254*[ id1_1]PX1
.constraint define 4 [ id1_1]PX1= 2*[id1_2_1]_cons
.gllamm wb1 PX1 PX2, nocons i(id) nrf(2) eqs(wgrp1 wgrp2) constraint(2 3 4) s(cons)
adapt nip(16)
```

The results are presented in Table 6. The same constraints were also applied using mvmeta and Stata ml. In this case, the agreement is up to the third decimal place, something expected if we consider the various transformations used. Nevertheless, for all practical applications, such agreement is more than sufficient, and the conclusions are not changed. After fitting the model, we can easily perform (using the delta-method) a formal comparison of the estimates in order to decide which treatment performs better:

```
.lincom _b[PX2]-_b[PX1]
```

The models presented in this section, being the more general (because they allow for a nonzero within-studies correlation), can be used without modifications for any other special cases of multivariate meta-analysis provided

| Table 5. Results for multiple-treatment meta-analysis for the sclerotherapy data. |
|------------------|------------------|------------------|------------------|------------------|
|                  | \( \beta_1 \) (SE) | \( \beta_2 \) (SE) | \( \tau_1^2 \) | \( \tau_2^2 \) | \( \rho \) |
| gllamm (method 1) | --                | --                | --              | --              | --              |
| gllamm (method 2) | -0.645007 (0.301752) | -0.59097 (0.278977) | 0.295302 | 1.021793 | 0.517541 |
| mvmeta (ML)       | -0.645007 (0.301746) | -0.59097 (0.278981) | 0.295301 | 1.021793 | 0.517538 |
| Stata ml          | -0.645007 (0.301751) | -0.59097 (0.278978) | 0.295301 | 1.021793 | 0.517540 |

Note that for this dataset, syntax 1 is not applicable.
ML, maximum likelihood; SE, standard error.
Table 6. Results for multiple-treatment meta-analysis for the sclerotherapy data assuming equal between-studies heterogeneity.

<table>
<thead>
<tr>
<th>Method</th>
<th>$\beta_1$ (SE)</th>
<th>$\beta_2$ (SE)</th>
<th>$\tau^2_1$</th>
<th>$\tau^2_2$</th>
<th>$\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td>gllamm (method 1)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>gllamm (method 2)</td>
<td>$-0.710884 (0.370972)$</td>
<td>$-0.584434 (0.256502)$</td>
<td>0.877349</td>
<td>0.877349</td>
<td>0.5</td>
</tr>
<tr>
<td>mvmeta (ML)</td>
<td>$-0.711140 (0.370947)$</td>
<td>$-0.584606 (0.256476)$</td>
<td>0.877160</td>
<td>0.877160</td>
<td>0.5</td>
</tr>
<tr>
<td>Stata ml</td>
<td>$-0.711140 (0.370948)$</td>
<td>$-0.584606 (0.256481)$</td>
<td>0.877160</td>
<td>0.877160</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Note that for this dataset, syntax 1 is not applicable.

ML, maximum likelihood; SE, standard error.

that this correlation can be calculated (Bagos, 2012; Wei and Higgins, 2012). For instance, one can easily fit models for bivariate meta-analysis of genetic association data (Bagos, 2008), models for surrogate markers (Daniels and Hughes, 1997), or more general multivariate models such as the trivariate model of (Arends et al., 2003). Of course, similar models can also be used, as we discussed earlier, for adjusting for the baseline risk by modeling simultaneously the log(odds ratio) and the baseline log(odds) as the pair of outcomes (Thompson et al., 1997; van Houwelingen and Senn, 1999). The advantage of such an approach compared with the one presented in the previous section is obtaining a direct estimate of the correlation and an easy Likelihood Ratio (LR) test for the relationship with the baseline risk, which is probably better than the delta method.

3.3. Multilevel meta-analysis and correlated estimates

Lastly, we are going to discuss a class of models for which no alternatives exist in Stata. Hence, we are now dealing with models with a single response variable (univariate), which, however, have observations that are clustered and/or correlated. In this class, we find the simple models for correlated estimates (Keller and Olkin, 2004), models for dose–response data (Berlin et al., 1993), models for meta-analysis of survival curves (Arends et al., 2008b), and multilevel models with their various instances (Thompson et al., 2001; Turner et al., 2000).

3.3.1. Correlated estimates. In this case, we have a single outcome in a model nearly identical to the simple univariate meta-analysis. The only difference is that the various estimates are not necessarily independent. This setting may arise, for instance, when some studies share a number of subjects (Gleser and Olkin, 1994; Keller and Olkin, 2004; Lin and Sullivan, 2009). In this case, the model is identical to that of Eq. (14):

$$y_i = \beta_0 + b_i + e_i, \quad b_i \sim N(0, \tau^2), \quad e_i \sim N(0, \Omega)$$  

with the exception that the within-studies covariance matrix is now given by

$$\Omega = \begin{bmatrix}
  \Sigma_1 & s_{12} & \cdots & s_{1k} \\
  s_{12} & \Sigma_2 & \cdots & s_{2k} \\
  \vdots & \vdots & \ddots & \vdots \\
  s_{1k} & s_{2k} & \cdots & \Sigma_k
\end{bmatrix}$$  

(36)

We should emphasize that there is no dedicated Stata command for fitting such models.

3.3.2. Multilevel models. A related but somehow more complex situation exists in the case of multilevel or hierarchical models (Thompson et al., 2001; Turner et al., 2000). Such cases may arise in situations where the estimates are nested. For instance, we may have estimates from different subgroups within the same study (Stevens and Taylor, 2009; Thompson et al., 2001), or estimates from studies that are grouped together based on some other criteria (Hemming et al., 2012).

In order to model such cases, we will now have subgroups (i) nested within studies (j) with their respective random terms:

$$y_{ij} = \beta + b_i + b_{ij} + e_{ij}$$  

$$b_i : N(0, \tau^2_i), \quad b_{ij} \sim N(0, \tau^2_j), \quad e_{ij} \sim N(0, \Omega)$$  

(37)

Here, $\tau^2_i$ is the between-studies variance, and $\tau^2_{ij}$ is the variance between groups within a study. In this model, the covariance matrix $\Omega$ will be of a block-diagonal form given by

$$\Omega = \text{diag}(\Sigma_1, \Sigma_2, \ldots, \Sigma_k) = \begin{bmatrix}
  \Sigma_1 & 0 & \cdots & 0 \\
  0 & \Sigma_2 & \cdots & 0 \\
  \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & \cdots & \Sigma_k
\end{bmatrix}$$  

(38)

where $\Sigma_i$ is the study-specific covariance matrix. There may be cases in which the estimates are independent within studies (Thompson et al., 2001), and thus, the study-specific covariance matrices will be of the form
In other cases, however, the estimates within studies are stochastically dependent similarly to what is discussed in the previous section (Stevens and Taylor, 2009), and thus, we will have

\[
\Sigma = \begin{bmatrix}
    s_{i1}^2 & 0 & 0 & \ldots \\
    0 & s_{i2}^2 & 0 & \ldots \\
    0 & 0 & s_{i3}^2 & \ldots \\
    \vdots & \vdots & \vdots & \ddots
\end{bmatrix}
\]  

(39)

In some cases, if the random variation between groups is found low, we can remove the respective term and model the random variation only in the study level:

\[
y_{ij} = \beta + b_i + \varepsilon_{ij}
\]

(40)

We will present the methods discussed in this section, using the data from a meta-analysis of 13 studies with 18 experiments all involving the effect of native-language (L1) vocabulary aids on second-language (L2) reading comprehension. Some studies produced multiple reports creating a dependence structure among the resulting effect size estimates (the authors use the Hedges's \(g\)). The covariance among these effect size estimates was estimated and incorporated into a proposed meta-analysis model that accounts for the dependence at a hierarchical level (Stevens and Taylor, 2009). In the beginning, we are going to disregard the grouping and model each subgroup as a separate study in order to mimic an analysis of data of the previous section. Afterwards, we are going to fit a truly multilevel model. The data are stored in the file `multilevel.dta`, and the necessary commands are given in the file `multilevel.do`. In order to fit a model that ignores both the hierarchical structure and the within-studies correlation, we would type

```stata
.gen logs=log(se)
.eq wgt: logs
.constraint define 1 [lns1]logs=1
.gllamm g, i(id) nrf(1) constraint(1) s(wgt) adapt nip(16)
```

whereas for fitting a hierarchical model that ignores only the within-studies correlation, we could use

```stata
.gllamm g, i(id study) nrf(1 1) constraint(1) s(wgt) adapt nip(16)
```

Note that these two models mentioned earlier ignore the within-studies correlation, and thus, they are inappropriate (they are presented only for illustration purposes). However, they can be fitted using syntax 1. For fitting model of Eq. (35), that is, treating each estimate as if it came from a different study, we will have to resort to syntax 2:

```stata
.mkmat v1-v18,mat(V)
.mat W=cholesky(invsym(V))
.mkmat g,mat(Y)
.mat wb=P*Y
.svmat wb
.gen cons=1
.mkmat cons,mat(C)
.mat PC=P*C
.svmat PC
.eq cons:cons
.eq slope:PC1
.constraint define 2 [lns1]cons=0
.gllamm wb1 PC1, nocons i(id) nrf(1) constraint(2) eq(slope) s(cons) adapt nip(16)
```

Finally, the multilevel model of Eq. (37) that takes into account both the hierarchical structure and the within-studies correlation is given in the following:
The analysis produces an estimate for mean $g$ equal to 0.698 (95% CI: 0.362, 1.034). The between-studies variance ($\tau^2_s$) is found equal to 0.269, whereas the between-groups variance ($\tau^2_g$) is equal to 0.027. These estimates are different but close to those reported by Stevens and Taylor (2009), which, however, used a different model and method of estimation. Because the between-groups variance ($\tau^2_g$) is quite low, it is tempting to use the more parsimonious model of Eq. (41):

$$y_{ij} = \beta x_{ij} + b_i x_{ij} + e_i, \quad b_i \sim N(0, \tau^2), \quad e_i \sim N(0, \Omega)$$  \hspace{1cm} (42)  

This model yields similar results, with a mean $g$ equal to 0.703 (95% CI: 0.373, 1.034) and a between-studies variance ($\tau^2_s$) equal to 0.293 (Stevens and Taylor, 2009).

3.3.3. Dose–response models. Another important univariate model with dependent errors is the model for meta-analysis of dose–response data. The data from such studies typically appear as a series of dose-specific relative risks, with one category serving as the common reference group; hence, such data exhibit within-studies correlation. Each category is associated with a specific dose (or a range of doses), and thus, we are interested in performing a regression in order to find the coefficient associated with a one-unit increase in dose. In the majority of such models, we also assume that a zero dose is associated with no risk, and thus, we fit a model with no intercept (Berlin et al., 1993; Greenland and Longnecker, 1992). Similar to the model of Eq. (37), we have subgroups ($j$) nested within studies ($i$):

$$y_{ij} = \beta x_{ij} + e_i, \quad e_i \sim N(0, \Omega)$$  \hspace{1cm} (38)  

with a covariance structure of the form of Eqs. (38) and (40). Here, the interest lies in the estimation of the regression coefficient, and the random effects are applied on the study level (random coefficient model). In Stata, there is a dedicated command, `glst` (Orsini et al., 2006), for fitting such models, but it uses only the method of moments. Here, we are going to fit the same model using `gllamm` and obtain the ML estimates. We will use the subset of the data from the meta-analysis for the effect of Body Mass Index (BMI) on the risk for renal cell cancer (Bergstrom et al., 2001), as reported by Liu et al. (2009). The data are given in the file `bmi-dose.dta`, and the respective commands in the do-file `dose-response.do`.

The analysis using the `glst` command yields a log(relative risk) of 0.019363 (95% CI: 0.013670, 0.025055) for fixed effects and 0.019312 (95% CI: 0.01065, 0.027978) for the random-effects analysis. Similarly, the analysis using `gllamm` yields an estimate of 0.019249 (95% CI: 0.013711, 0.024788) for the fixed effects and 0.019087 (95% CI: 0.011245, 0.026928) for the random-effects analysis. These estimates are very close, despite the different method of estimation. The between-studies variance is small, and it was estimated equal to 0.0000572 with `gllamm` and 0.0000807 with `glst`.  

$$y_{ij} = \beta x_{ij} + e_i, \quad e_i \sim N(0, \Omega)$$  \hspace{1cm} (38)
Of course, *gllamm* being very flexible allows for several extensions as well. For instance, we can perform a two-stage approach pooling first the estimates within studies and then combining them with traditional methods, whereas the model itself can be modified to include nonlinear terms (Orsini et al., 2012), or nonzero intercept (Smith et al., 1995). As a matter of fact, the two-step approach does not even require *gllamm*, because it can also be carried out using *glm* and *metan* as follows:

```
. gen b_est=.
. gen se_est=.
. sort id
. by id: gen st =id if _n==_N
. local N=`_N'
. forvalues x=1(1)`N' {
. glm wbl PX1 if id==`x', nocons scale(1)
. mat b =e(b)
. mat V =e(V)
. replace b_est=b[1,1] if st==`x'
. replace se_est=sqrt(V[1,1]) if st==`x'
}
. metan b_est se_est, randomi
```

Finally, there are several other models with a univariate response and correlated errors that can be very easily fitted using the method outlined in the preceding text. These include the method of Arends and coworkers for meta-analysis of survival proportions reported at multiple time points (2008b) and methods for meta-analysis of repeated measures and longitudinal data (Jones et al., 2009; Ishak et al., 2007; Peters and Mengersen, 2008).

### 4. Discussion

In this work, we showed with many practical examples, which many seemingly unrelated models, including univariate meta-analysis and meta-regression, multivariate meta-analysis, meta-analysis of dose-response models, meta-analysis with correlated estimates, multilevel meta-analysis, and meta-analysis of longitudinal data, can be fit using *gllamm*. The software is very versatile and can handle a wide variety of models with applications in a wide range of disciplines. Here, not only were we able to replicate the analyses performed by some dedicated *Stata* commands (such as *metan*, *metareg*, *mvmeta*, and *glst*) but we also showed how to fit some other models for which alternative commands do not exist. The method presented here takes advantage of the modeling capabilities of *gllamm* and its flexibility and makes use of appropriate transformations based on the Cholesky decomposition of the inverse of the covariance matrix, known as GLS, in order to handle correlated data. All the models described in the preceding text can be thought of as special instances of a linear mixed-model formulation that has been proposed in the past (Platt et al., 1999; Stram, 1996), but to the author’s knowledge, a general exposition in order to incorporate all the available models for meta-analysis as special cases and the instructions to fit them in *Stata* has not been presented so far. Similar models can be fitted in *SAS* using PROC MIXED, following the excellent tutorial of van Houwelingen et al. (2002). The same models can also be fitted using the *lme()* function in *S-PLUS*, Insightful Corporation, Seattle, WA, but not in *R* (see the discussion in http://www.metafor-project.org/doku.php/tips:rma_vs_lm_and_lme). Instead, *R* users should use the *metafor* package (Viechtbauer, 2010).

As we already noticed, the present work deals exclusively with meta-analysis of summary-based data, for which the GLS approach is appropriate. *gllamm* can also be used in meta-analysis, taking advantage of its capabilities in estimating generalized linear mixed models that directly use binary or count data. Such methods, in different contexts, can be termed IPD methods, or models based on the exact likelihood, but we need to emphasize that usually, IPD are not available and these methods in the majority of the situations just use the counts extracted from the published reports. In any case, when applicable, such methods have important advantages compared with the usual summary-data methods that rely on the normal approximation because they avoid the normality assumptions and they do not need continuity correction in case of zero cell counts, resulting thus in unbiased estimates (Hamza et al., 2008).

Models for binary data that use the odds ratio have been presented earlier (van Houwelingen et al., 1993; Turner et al., 2000). A tutorial for such meta-analysis in *gllamm* can be found in the work by Skrondal and Rabe-Hesketh (2004) (the code is available at http://www.gllamm.org/books/gum.html). Some important special cases include the meta-analysis of diagnostic tests (Chu and Cole, 2006) and the meta-analysis of genetic association studies (Bagos, 2008; Bagos and Nikolopoulos, 2007) (for the code, see http://www.stata.com/statalist/archive/2004-04/msg00820.html and http://www.compagen.org/tools/multivariate-genetic, respectively).
Other models, based on the Poisson likelihood, have been presented for analyzing studies with varying lengths of follow-up (Bagos and Nikolopoulos, 2009) (the code is available at http://www.compgen.org/tools/poisson-metanalysis). Models using the multinomial likelihood have been proposed for meta-analysis of genetic association studies (Bagos, 2008) (http://www.compgen.org/tools/multivariate-genetic), whereas methods for meta-analysis of studies with ordinal outcomes have also been proposed (Bipat and Zwinderman, 2010; Whitehead et al., 2001; Whitehead and Jones, 1994). To the author's knowledge, there is no application of gllamm in meta-analysis of ordinal outcomes, but the functionality of the tool allows, at least, some of the aforementioned models to be fitted. Models using the binomial or the Poisson likelihood can also be fitted in SAS using PROC NLMIXED. The interested reader may refer to the excellent tutorial by Stijnen and coworkers, which presented several examples and made the source code available (2010).

Despite the advantages mentioned earlier, exact likelihood methods suffer from some disadvantages as well. First of all, they are in general more computational demanding (note, however, that this is not the case for models fitted with gllamm because the program uses the same algorithms for likelihood maximization). Secondly, they are applicable only in selected cases. As a matter of fact, few of the methods presented in this work can be fitted using the exact likelihood (i.e., the univariate meta-analysis and the bivariate meta-analysis for multiple treatments). On the other hand, earlier works have shown that, at least under certain circumstances, the summary methods are equivalent to the IPD methods (Olkin and Sampson, 1998, Mathew and Nordstrom, 1999). These results were generalized recently, and theoretical evidence suggests that using time-consuming individual data methods for meta-analysis of large studies does not necessarily lead to increased efficiency (Lin and Zeng, 2010a, 2010b). Empirical findings also corroborate these conclusions, suggesting that in practical applications, summary-based methods and IPD perform similarly (Janssens et al., 2009).

As we already mentioned, gllamm uses adaptive quadrature, which is an approximate method (Rabe-Hesketh et al., 2002). Thus, one may argue that its use in cases of normally distributed responses may not be acceptable and other tools that directly model the likelihood would be more appropriate. However, there is a large body of literature with successful and reliable applications of gllamm in such models (Rabe-Hesketh and Skrondal, 2004). Similar observations hold also for all the models tested in this work. In all cases where reliable software of ML estimation exists, gllamm performs nearly identical, with the possible downside of increased computational time. On the other side, the flexibility of gllamm is unique, because it allows fitting many diverse models, provided that the model can be written in the quite general form of Eq. (1). Although for some of these models, reliable alternatives do exist, it is of importance to notice, once again, that hierarchical models for meta-analysis as well as models for repeated measures and longitudinal data can be fitted only using gllamm. Thus, we think that the approach outlined here is useful in many respects and serves a manifold role. Besides the obvious advantage of showing in practice how the various models can be fitted in Stata, the manuscript, integrating all these models in a unified framework, serves also a pedagogical role. Additionally, using a regression framework allows the full use of the relevant machinery, that is, the ability to check modeling assumptions, to detect outliers, or to perform formal model comparisons. All data and source code needed for these analyses is given in www.compgen.org/tools/gllamm, and we hope that they will be used in practical applications.

**Appendix**

**gllamm** is not a standard module in Stata, but it is freely available, and the installation is straightforward. There is a dedicated webpage, created by the authors of the software, where tutorials, instructions, examples, and references are listed (www.gllamm.org), and the user may refer to this page for more details. The easiest way of installing the current version of the program is to use the ssc commands from within Stata (help ssc):

```
.ssc describe gllamm
.ssc install gllamm
```

Additional details can be found on http://www.gllamm.org/install.html.

In the same page, a list with various tutorials, published worked examples, and papers that used gllamm is available (http://www.gllamm.org/examples.html). The beginner may start with the excellent tutorial by Stas Kolenikov http://web.missouri.edu/~kolenikovs/stata/gllamm-demo.html and the review by Grilli and Rampichini (2006), which as a matter of fact presents an example for univariate meta-analysis. Moreover, the readers may find many worked examples in the book written by gllamm authors (Skrondal and Rabe-Hesketh, 2004).

For other dedicated routines for meta-analysis in Stata, which were mentioned earlier, the interested user should consult the respective publication in Stata Journal. A non-inclusive list of the available commands can be found on http://www.stata.com/support/faqs/statistics/meta-analysis/. These commands are discussed in the excellent tutorial of Sterne and coworkers (2001).
Acknowledgements

This work was funded by the project "IntDaMuS: Integration of Data from Multiple Sources," which is implemented under the "ARISTEIA II" Action of the "Operational Programme Education and Lifelong Learning" and is co-funded by the European Social Fund (ESF) and National Resources. The author would like to thank the editor in chief, the associate editor, and the two anonymous reviewers for the constructive comments that helped in improving the quality of the manuscript. Maria Adam and Orestis Efthimiou are also acknowledged for the fruitful discussions that improved the manuscript.

References


