Statistical Models and Methods for Network Meta-Analysis

L. V. Madden, H.-P. Piepho, and P. A. Paul

First and third authors: Department of Plant Pathology, Ohio State University, Wooster 44691; second author: Biostatistics Unit, Institute of Crop Science, University of Hohenheim, 70599 Stuttgart, Germany.
Accepted for publication 16 April 2016.

ABSTRACT

Madden, L. V., Piepho, H.-P., and Paul, P. A. 2016. Statistical models and methods for network meta-analysis. *Phytopathology* 106:792-806.

Meta-analysis, the methodology for analyzing the results from multiple independent studies, has grown tremendously in popularity over the last four decades. Although most meta-analyses involve a single effect size (summary result, such as a treatment difference) from each study, there are often multiple treatments of interest across the network of studies in the analysis. Multi-treatment (or network) meta-analysis can be used for simultaneously analyzing the results from all the treatments. However, the methodology is considerably more complicated than for the analysis of a single effect size, and there have not been adequate explanations of the approach for agricultural investigations. We review the methods and models for conducting a network meta-analysis based on frequentist statistical principles, and demonstrate the procedures using a published multi-treatment plant pathology data set. A major advantage of network meta-analysis is that correlations of estimated treatment effects are automatically taken into account when an appropriate model is used. Moreover, treatment comparisons may be possible in a network meta-analysis that are not possible in a single study because all treatments of interest may not be included in any given study. We review several models that consider the study effect as either fixed or random, and show how to interpret model-fitting output. We further show how to model the effect of moderator variables (study-level characteristics) on treatment effects, and present one approach to test for the consistency of treatment effects across the network. Online supplemental files give explanations on fitting the network meta-analytical models using SAS.

Additional keywords: Fusarium head blight of wheat, linear mixed models, mixed treatment comparisons, multiplicative interactions.

Meta-analysis is very common in many disciplines for combining results from multiple studies (Borenstein et al. 2009; Higgins et al. 2009; Lipsey and Wilson 2001). The popularity of the approach has grown enormously over the past 4 decades (Madden and Paul 2011). Based on the Web of Knowledge, there are now over 98,000 published articles where meta-analysis is used or the methodology is developed, with over 14,000 articles in 2015 alone. After a slow start, there is increasing use of meta-analysis for quantitative research synthesis in plant pathology and other agricultural fields (Madden and Paul 2011; Makowski et al. 2014; Ngugi et al. 2011a; Philibert et al. 2012; Scherm et al. 2014). Most meta-analyses involve a single summary result (the so-called estimated effect size) from each study, such as the mean, difference of means for two treatments, standardized difference of means for two treatments, log response ratio, log odds, and log odds ratio (Borenstein et al.

2009; Ngugi et al. 2011a; Paul et al. 2007, 2011). However, it is not uncommon for there to be multiple effect sizes per study. This may correspond to the situation where there are (i) two or more response variables for each study, such as the estimated slope and intercept for a relationship between two variables (e.g., disease severity and yield loss, or disease severity and toxin concentration of grain [Paul et al. 2006; Riley 2009; Salgado et al. 2015]); or (ii) where there are two or more treatments per study (Jackson and Riley 2014; Jackson et al. 2013; Riley et al. 2007). This review is concerned with the multi-treatment problem.

When there are multiple treatments, researchers often perform a separate meta-analysis for each effect size of interest (e.g., for each pair of treatment means). Such analyses ignore the correlation of effect size estimates within studies, which could lead to biased overall results (Higgins et al. 2012; Jones et al. 2011). Moreover, different collections of studies may be used for different pairwise mean differences, because all treatments typically are not in all available studies. This means that the evidence base is different for each effect size of interest (Tu 2014, 2015).

There are many excellent articles and books describing the principles and practice of meta-analysis for single effect sizes per study (Borenstein et al. 2009; Higgins et al. 2009; Lipsey and Wilson 2001; Whitehead 2002). Madden and Paul (2011) recently

Corresponding author: L. V. Madden; E-mail address: madden.1@osu.edu

*The e-Xtra logo stands for "electronic extra" and indicates that one supplementary appendix and two supplementary SAS files are published online.

<http://dx.doi.org/10.1094/PHYTO-12-15-0342-RVW>
© 2016 The American Phytopathological Society

presented a detailed synthesis of the approach for problems in plant pathology. There are far fewer reviews of multi-treatment meta-analysis, even though such analyses are now becoming more common in many disciplines (Higgins et al. 2012; Jones et al. 2011; Salanti 2012; Salanti et al. 2008), including in plant pathology (Madden et al. 2014; Nelson et al. 2015; Ngugi et al. 2011b; Paul et al. 2008, 2010; Willyerd et al. 2012). The purpose of this paper is to review the concepts and methods for multi-treatment meta-analysis, and describe models that are appropriate for conducting the analysis. We cover relevant approaches that have been developed for analyzing results from multiple studies in agriculture, medicine, or elsewhere. A previously published multi-treatment data set (Paul et al. 2010) is used to demonstrate some of the different analyses that are possible.

OVERVIEW OF MULTI-TREATMENT META-ANALYSIS

Suppose a collection of individual studies contains treatments A, B, and C, with C being the control. In separate univariate analyses, one could consider each pairwise treatment comparison of interest (e.g., d_{AC} , d_{BC} , d_{AB} , in which d is the difference of treatment means in a study) in separate univariate analyses. This approach, however, does not take into account the correlations of the pairwise differences; these correlations can affect the results for any treatment comparison. As an alternative, one can be more efficient and conduct an overall multi-treatment analysis simultaneously involving all the treatments of possible interest (Higgins et al. 2012; Salanti et al. 2008; Whitehead 2002). This allows direct comparisons of all the treatments with each other, and takes into account all the correlations.

The multi-treatment analysis can also use more of the individual studies. For instance, if the primary interest was the A-B mean difference, and one was conducting a typical univariate meta-analysis of d_{AB} , one could only use studies where both treatment A and B were present. However, with a multi-treatment analysis, one can use studies with either one of the treatments being present; for instance, the lack of treatment B in a given study does not preclude the use of studies that contain only A and C, or the lack of A does not preclude use of studies with B and C only. Studies with both A and B give direct evidence of the difference between the two means, whereas studies with only A or B (but both containing C) give indirect evidence of the difference between the A and B means (because $d_{AB} = d_{AC} - d_{BC}$) (Higgins et al. 2012; Jones et al. 2011, Lu and Ades 2006).

Several labels have been given to meta-analysis with multiple treatments per study (Salanti 2012). Lu and Ades (2004) coined the term mixed treatment comparisons, and earlier Lumley (2002) used the term network meta-analysis. Network meta-analysis is probably the most common label in use, especially in the medical area (Salanti and Schmid 2012). We consider these synonyms, and primarily use the term network meta-analysis throughout the paper.

Several approaches have been proposed and utilized for network meta-analysis, mostly based on linear mixed models (Bagos 2008; Greco et al. 2015; Higgins et al. 2012; Jones et al. 2011; Lu and Ades 2004, 2006; Salanti et al. 2008; Tu 2014; van Houwelingen et al. 2002; Whitehead 2002). The Bayesian approach of Lu and Ades has become especially common over the last few years (Higgins et al. 2012; Salanti and Schmid 2012). If the original replicate data from each study are available, a single-stage analysis can be performed, using the original observations from the individual studies (Whitehead 2002). Basically, one utilizes a hierarchical linear or generalized linear mixed model, similar to what is done for multicenter clinical studies (Brown and Prescott 2006) or multi-environment plant variety trials (Denis et al. 1997; Piepho 1999; Smith et al. 2005). Riley et al. (2010) give several advantages to the individual-observation single-stage analysis. It is most advantageous when the individual studies are small (e.g., with few replications). This type of meta-analysis has recently been done in plant pathology (Nelson et al. 2015). More commonly,

only the summary results (e.g., estimated effect sizes [means], and associated standard errors or variances) for each treatment may be available from individual studies, requiring a two-stage analysis. The first stage entails the analysis of each individual study; this is often done by the investigators conducting the individual primary studies, although this can also be done by the meta-analyst. Then, meta-analysis is performed on summary results in the second stage using a mixed model.

Piepho et al. (2012a) and Möhring and Piepho (2009) show that the two-stage approach, based on the proper model formulation, often gives very similar results compared with those obtained with the single-stage approach. Moreover, it may not be practical to do a single-stage analysis if the experimental and treatment designs vary among the studies; for instance, different studies may be split plots, randomized complete blocks, or repeated measures. This diversity of study types is not an obstacle with the typical two-stage approach, as long as one is careful in extracting the summary results and determining the proper sampling variances to use in the second-stage (Borenstein et al. 2009). This is discussed more in the appendix in the on-line supplemental file. It should be pointed out that the work by Piepho et al. (2012a) and Möhring and Piepho (2009) was motivated by computational demands; they were concerned with large numbers of plant cultivars ("treatments") per study (environment), making it difficult or too time consuming to fit a mixed model to the observations from all studies in one step. Computer memory may not permit a single-stage analysis in these situations. With meta-analysis, the motivation is different: the data analyst typically does not have access to all the original observations. Yet the findings by Piepho et al. (2012a) about two-stage analyses are relevant for both types of problems. This paper focuses on the two-stage meta-analysis.

The effect sizes used in a network meta-analysis can be of two different forms as explained below (Salanti et al. 2008; Salanti 2012). Most methods operate on pairwise differences (contrasts; e.g., d_{AB}) of treatment means (Glesser and Olkin 2009; Higgins et al. 2012; Krahn et al. 2013; Lu and Ades 2004; White et al. 2012). That is, for each study, one constructs contrasts of each treatment with a so-called baseline treatment, which is usually the control. Then one analyzes jointly the contrasts across all the studies. We call this the conditional modeling approach (Piepho et al. 2012b). The contrast estimates are correlated within studies, by definition, because they involve a common baseline; thus, the within-study correlations (or covariances) and variances should be accounted for in the meta-analysis. Setting up the contrasts and variance-covariance matrix of estimated contrasts for each study can be challenging for this network approach (Higgins et al. 2012; Lu and Ades 2006). This is especially challenging when the baseline treatment is not the same for all studies. For instance, if there were four treatments (A, B, C, D, with C being considered the baseline), but a study was conducted without C, then another treatment must be considered the baseline for that study in order to obtain the contrasts. This, in turn, requires considerable care in setting up the linear mixed model indicator variables (dummy variables) for that study in order to obtain meaningful results (Piepho et al. 2012b; Piepho 2014). There is considerable danger that the analyst could make mistakes in the model formulation or in the construction of the covariance matrix.

The implementation of the Lu and Ades approach has received considerable attention recently, with advice on setting up the contrasts, variance-covariance matrices, and mixed-model variables (Dias et al. 2013a,b; Donegan et al. 2013; Higgins et al. 2012; Hoaglin et al. 2011; Jones et al. 2011; Lu et al. 2011; Ohlssen et al. 2014; Salanti et al. 2008; Tu 2014). A special issue of *Research Synthesis Methods* focuses on many aspects of this baseline-contrast approach to network meta-analysis (Salanti and Schmid 2012). With some exceptions (Willyerd et al. 2012), plant pathologists have not taken this approach to meta-analysis involving multiple treatments.

As an alternative for network meta-analysis, a two-way linear mixed model can be fitted directly to the treatment means from each study in a two-stage analysis (Greco et al. 2015; Piepho 2014; Piepho et al. 2015; van Houwelingen et al. 2002). We call this the unconditional modeling approach. When the original observations are present, one can also use this second modeling approach to analyze the observations directly without defining contrasts with a baseline treatment (Greco et al. 2015; Piepho et al. 2012b; Whitehead 2002). The two modeling approaches are closely related, and Piepho et al. (2012b) recently showed that they are equivalent under certain conditions. The two-way unconditional linear mixed model is much easier to implement with standard software than the conditional model (Greco et al. 2015). Only treatment summaries such as means (and corresponding measures of within-study variability) are needed for the two-stage approach (Piepho 2014; Piepho et al. 2012b).

The two-way unconditional linear mixed model is commonly used in plant pathology for multi-treatment meta-analysis (Madden et al. 2014; Paul et al. 2008, 2010; Nelson et al. 2015; Ngugi et al. 2011b). However, several versions of the model can be used in the analysis (Bagos 2008; Brown and Prescott 2006; Piepho 2014; Piepho et al. 2015; Senn 2010; Senn et al. 2013; van Houwelingen et al. 2002; Whitehead 2002). A major decision of the analyst is whether the study main effect should be considered fixed or random; the implications of this choice are discussed below in two different sections. Although detailed methods have been published for conducting the contrast-based conditional modeling approach of Lu and Ades (Dias et al. 2013a; Jones et al. 2011; Krahn et al. 2013; Lu et al. 2011; Ohlssen et al. 2014; Tu 2014; White et al. 2012), much less attention has been given to the unconditional modeling approach. Below we present and compare several of the unconditional models that are applicable for a meta-analysis, with an emphasis on the second-stage analysis, where the treatment summaries from the first stage of the analysis comprise the available data set. We also present some comparisons with the contrast-based approach with summary data. Finally, the method described in Piepho (2014) and Piepho et al. (2012b, 2015) for determining the consistency of treatment effects in a network meta-analysis is discussed. An example data set is utilized to demonstrate the different analyses.

EXAMPLE

Paul et al. (2010) reported on a meta-analysis of wheat yield (metric tons per hectare) in field plots that were treated with different fungicides to control Fusarium head blight. A total of 136 studies—arising from investigations in 14 different U.S. states over 12 years—were analyzed. Each study was laid out as a randomized complete block design, although the number of blocks varied among the individual studies. There were a total of six treatments, including the no-fungicide control. All studies had a no-fungicide control and a tebuconazole (Folicur) treatment, and from zero to four additional treatments (propiconazole [Tilt], prothioconazole [Proline], tebuconazole + prothioconazole [Prosaro], and metconazole [Caramba]). The average number of treatments per study was four. From the first-stage analysis, based on an analysis of variance (ANOVA) for yield, the mean yield (and corresponding sampling variance) for each treatment of each study was determined; these means were used as the summary results in the second-stage analyses considered below. The means are assumed to be normally distributed. The sampling variance used for each mean within a study was one-half of the squared standard error of the difference (SED) of the means (for balanced data, this is the residual variance divided by the number of blocks) (Madden and Paul 2011). Online Supplementary appendix S discusses the appropriate sampling variance for different within-study experimental designs. The full data set is available in the online SAS supplementary files (Supplementary Files S1 and S2).

MIXED MODELS

Basic model. The summary result (or estimated effect size) for the i th study ($i = 1, \dots, K$) and j th treatment ($j = 1, \dots, m$) is given by y_{ij} . For our example, the estimated effect size is the mean, that is, the estimated expected value for a given treatment. The summary result could also be a transformation of the mean (Paul et al. 2008) or other statistic. The number of treatments can vary among studies.

The estimated sampling variance of y_{ij} within the i th study is given by s_{ij}^2 , although it is possible that this variance is the same for all treatments in a study (so that $s_{ij}^2 = s_i^2$). The latter occurs when a normal distribution with variance homogeneity across treatments is assumed for the observations within a study and a simple linear (mixed) model is fitted to the individual study data. As is the common practice, we assume for most of the subsequent analyses that the means for the different treatments within a study are independent (i.e., that the means conditional on the study are independent). The means are not actually independent with randomized complete block designs (when block is considered a random effect), but use of one-half of the square of the SED for s_{ij}^2 permits the use of this simplifying assumption for this design (see Supplementary appendix S and Möhring and Piepho 2009).

A model for the summary results is given by

$$y_{ij} = \eta_{ij} + e_{ij} \quad (1)$$

where η_{ij} is the so-called linear predictor for the effects of the i th study and j th treatment, and e_{ij} is the residual, which represents the error in estimating the summary result (effect size) for treatment j of study i . It is assumed that e_{ij} is normally distributed with a mean of 0 and variance of s_{ij}^2 ; that is, $e_{ij} \sim N(0, s_{ij}^2)$. It is convenient to combine the effects associated with the treatments in a study into vectors. In matrix notation, for the i th study, one can define $\mathbf{y}_i$ as the vector of the summary results, $\boldsymbol{\eta}_i$ as the vector of linear predictors, and $\mathbf{e}_i$ as the vector of residuals. For ease of expression, we generally write out the vectors and matrices for studies with all m treatments, although a study can have any number of treatments equal to or less than m . For instance, we can write

$$\mathbf{y}_i = (y_{i1}, y_{i2}, \dots, y_{im})^T, \quad \boldsymbol{\eta}_i = (\eta_{i1}, \eta_{i2}, \dots, \eta_{im})^T, \\ \mathbf{e}_i = (e_{i1}, e_{i2}, \dots, e_{im})^T$$

The matrix version of equation 1 is

$$\mathbf{y}_i = \boldsymbol{\eta}_i + \mathbf{e}_i \quad (2)$$

in which $\mathbf{e}_i \sim N(\mathbf{0}, \mathbf{R}_i)$, where $\mathbf{R}_i$ is the variance-covariance matrix for the residuals [$\text{var}(\mathbf{e}_i) = \mathbf{R}_i$] for the i th study. For the independent treatment situation considered here, one can write

$$\mathbf{R}_i = \begin{pmatrix} s_{i1}^2 & 0 & \cdots & 0 \\ 0 & s_{i2}^2 & & \vdots \\ \vdots & & \ddots & \\ 0 & & & s_{im}^2 \end{pmatrix} \quad (3)$$

when the study contains all m treatments. In other situations $\mathbf{R}_i$ may not be a diagonal matrix. In meta-analytical model fitting, the residual variances (or $\mathbf{R}_i$) for each study are held fixed (Higgins et al. 2009; Madden and Paul 2011). This requires care in preparing code for statistical programs.

A key issue in data analysis is the choice for the form of η_{ij} . Major possibilities are considered in the next sections that are based on a linear mixed model with study and treatment main effects. Appendix A provides a guide for using the GLIMMIX procedure in SAS to fit the models.

LINEAR MODELS WITH FIXED STUDY EFFECTS

Homogeneity of among-study variance. For the basic fixed-study-effects model, one can write the linear predictor as

$$\eta_{ij} = \theta + \beta_i + \tau_j + u_{ij} \quad (4)$$

in which θ is a constant (intercept), β_i is the fixed main effect of the i th study, τ_j is the fixed main effect of the j th treatment, and u_{ij} is the random effect associated with the linear predictor, which represents the between-study variability in the treatment effect. For identifiability, there are constraints on the fixed effects. One common constraint is to force the parameter for the last level of an effect (e.g., τ_m) to be 0. The linear predictor η_{ij} gives the expected value of y_{ij} conditional on any nonresidual random effects. The u_{ij} effect is equivalent to a random study $\times$ treatment interaction. The simplest assumption for the random effect is that it has a normal distribution with mean 0 and constant variance σ_u^2 ; that is, $u_{ij} \sim N(0, \sigma_u^2)$. We can define $\mathbf{u}_i$ as the vector of between-study effects for the i th study, and write $\mathbf{u}_i \sim N(\mathbf{0}, \Sigma)$, where Σ is the variance-covariance matrix of the between-study effects. The matrix is given in Table 1; for this model, the between-study covariance (or correlation) of the random effect is 0. This model, with between-study variance homogeneity, is labeled model F1.

Combining equations 1 and 4, one obtains for model F1:

$$y_{ij} = \theta + \beta_i + \tau_j + u_{ij} + e_{ij} \quad (5)$$

The expected value for treatment j across all the studies is given by

$$\mu_j = \theta + \bar{\beta}_\bullet + \tau_j$$

where

$$\bar{\beta}_\bullet = \sum_i^K \beta_i / K$$

is the mean study effect. The difference in expected value for any pair of treatments (j, j') is given by $\mu_j - \mu_{j'} = \tau_j - \tau_{j'}$. Piepho et al. (2012b) showed that fitting this linear mixed model (i.e., with a fixed study main effect) to the summary data using restricted maximum likelihood (REML) is equivalent to fitting the baseline-contrast conditional model of Lu and Ades (2004). Identical results are obtained with both approaches (see also Piepho 2014). In Appendix B, we demonstrate the direct link between model F1 and the baseline-contrast model. When analyzing replicate data in a one-stage analysis, the results of the two approaches—with or without baseline contrasts—are similar but not identical.

Model F1 is analogous to the model for an incomplete block design with fixed block effects. Unlike with an incomplete block design, there is no randomization of individuals (plots, plants, etc.) among the studies, only within the studies. Moreover, the presence or absence of a given treatment in a study typically is also not under the control of the meta-analyst (nor is it typically determined by a randomization process). With the use of equation 4, in which the study main effect is assumed to be fixed, treatment effects (e.g., $\tau_1, \tau_1 - \tau_2$) are estimated just based on intrastudy information (Piepho et al. 2012b; Riley et al. 2008; Senn 2010; Senn et al. 2013). Senn (2010) considers it to be desirable to use only intrastudy information because the principle of concurrent control is respected. When there is concurrent control, treated and untreated wheat plots (in our example) come from the same source of plants, and, in principle, wheat seeds are selected randomly to be in the treated or control plots. Several authors have advocated the use of study as a fixed effect for meta-analysis (Riley et al. 2008). Note, that variation in

treatment effects among studies is still explicitly utilized with this model through the u_{ij} term.

Generally, with network meta-analysis, a compilation of studies with different sets of treatment gives rise to an incomplete treatment-by-study classification. That is, there will be from a few to many empty cells in a study $\times$ treatment table. All treatment comparisons are estimable using model F1 (and the following models) in such a setting, provided that the treatment-by-study classification is connected (Searle 1987). Piepho et al. (2012b) discusses at length the meaning of connectedness in this context. The classification is said to be connected if each pair of filled cells in the study $\times$ treatment table can be connected by a continuous line, consisting solely of horizontal and vertical segments, and having changes of direction only in filled cells (Searle 1987, page 140).

Fitting model F1 results in estimates of the parameters, in particular the treatment effects ($\hat{\tau}_j$) and the estimated variances of the estimated effects [$\text{var}(\hat{\tau}_j)$]. Generally, differences involving two treatments (e.g., $\hat{\tau}_1 - \hat{\tau}_2$, equivalent to $\hat{\mu}_1 - \hat{\mu}_2$), are of primary interest with multi-treatment meta-analysis. Because of the unequal residual (sampling) variances across the studies (and possibly unequal residual variances across treatments within studies), the estimated variance of the difference [$\text{var}(\hat{\tau}_j - \hat{\tau}_{j'})$; e.g., $\text{var}(\hat{\tau}_1 - \hat{\tau}_2)$] has a complex formula that depends on the variances of u_{ij} and e_{ij} (Littell et al. 2006); mixed-model software automatically determine these variance estimates. For intuitive insight, it is useful to point out that the portion of $\text{var}(\hat{\tau}_j - \hat{\tau}_{j'})$ based on the between-study variability equals $\hat{\sigma}_u^2 + \hat{\sigma}_u^2 = 2\hat{\sigma}_u^2$ (when equation 4 is fitted). The between-study portion of the estimated covariance between two contrasts involving a common treatment [e.g., $\text{cov}(\hat{\tau}_1 - \hat{\tau}_2, \hat{\tau}_1 - \hat{\tau}_3)$] equals $\hat{\sigma}_u^2$.

When there are only two treatments per study and independent treatments within the studies, then fitting model F1 is equivalent to fitting a random-effects model to the result summaries of $D_i = y_{i1} - y_{i2}$, where the within-study (residual) variance is $s_{i1}^2 + s_{i2}^2$ (held fixed), and the between-study variance is $2\hat{\sigma}_u^2 (= \hat{\sigma}_D^2)$ (Madden and Paul 2011).

SAS GLIMMIX code for fitting model F1 (and other models) using REML is given in Appendix A and in the SAS program file (Supplementary Files S1 and S2). More detailed explanations are given in Piepho et al. (2012b) and Piepho (2014).

Variance heterogeneity. Several generalizations of model F1 are possible when study is a fixed effect. One natural extension of equation 4 is to allow for unequal variances for the between-study heterogeneity, so that each treatment has a different variance ($\sigma_{u(j)}^2$). The linear predictor has the same form (equation 4), but one writes $u_{ij} \sim N(0, \sigma_{u(j)}^2)$. The variance-covariance matrix of $\mathbf{u}_i$ (Σ) for this model is given in Table 1; there also is zero between-study covariance of the u_{ij} values with this situation. We refer to this as model F2. The portion of $\text{var}(\hat{\tau}_j - \hat{\tau}_{j'})$ based on the between-study

TABLE 1
Between-study variance-covariance matrix for meta-analytical models with a fixed effect of study

Model ^a	$\Sigma = \text{var}(\mathbf{u}_i)$
F1	$\begin{pmatrix} \sigma_u^2 & 0 & \cdots & 0 \\ 0 & \sigma_u^2 & & \\ \vdots & & \ddots & \\ 0 & & & \sigma_u^2 \end{pmatrix}$
F2	$\begin{pmatrix} \sigma_{u(1)}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{u(2)}^2 & & \\ \vdots & & \ddots & \\ 0 & & & \sigma_{u(m)}^2 \end{pmatrix}$

^a Model F3 has the same Σ structure as model F1.

variability now equals $\hat{\sigma}_{u(1)}^2 + \hat{\sigma}_{u(2)}^2$. The between-study portion of the covariance formula for two contrasts involving a common treatment (e.g., $\text{cov}(\hat{\tau}_1 - \hat{\tau}_2, \hat{\tau}_1 - \hat{\tau}_3)$) equals $\hat{\sigma}_{u(1)}^2$ when treatment 1 is the common treatment.

More complex structures for the Σ variance-covariance matrix can be considered. For instance, nonzero correlations of the u_{ij} effects in equation 4 can be accounted for using a so-called factor-analytic structure (Piepho 1997). We do not pursue this generalization here, but readers may find a multi-treatment meta-analysis example in Piepho et al. (2012b, page 1272).

Example. Based on the AIC values and a likelihood ratio test (difference of -2 times the log-likelihoods = 29.1, $\text{df} = 5$, $P < 0.001$), model F2 provided a better fit to the data than did model F1 (Table 2). Seven of the fifteen pairwise contrasts of the six means are shown in Table 3. Between-study variance estimates for model F2 ranged from 0 to 0.062 for the six treatment means (with the largest variance for the control). Nevertheless, the estimates of the pairwise contrasts and the SEs of these estimated contrasts were quite similar for the two models with this data set, with the range of SEs being slightly larger for model F2 than for F1 (Table 3).

MULTIPLICATIVE MODELS WITH FIXED STUDY EFFECTS

Multiplicative models involving study and treatment as fixed effects have been proposed by different authors, primarily for the two-treatment problem (i.e., treatment and a control) (Arends et al. 2003; Schmid et al. 1998; Sharp and Thompson 2000). Piepho et al. (2015) recently proposed this approach for the general multi-treatment meta-analysis. Ignoring heterogeneity among the studies (i.e., dropping the u_{ij} term), the expected treatment difference for a pair of treatments can be written as

$$\eta_{i2} - \eta_{i1} = \vartheta + \xi \eta_{i1} \quad (6)$$

where the linear predictor η_{i1} is the expected value of the control treatment in the i th study, and η_{i2} is the expected value of the other (new) treatment. The parameter ξ gives the sensitivity of the treatment effect (i.e., difference of treatment means) to the magnitude of the

expected value of the control; ϑ is the expected value of the treatment effect (difference of treatment means) when the expected value of the control is 0. When $\xi < 0$, the treatment effect declines as the control mean increases; when $\xi > 0$, the treatment effect increases with the control mean. Equation 6 is equivalent to $\eta_{i2} = \vartheta + (\xi + 1)\eta_{i1}$, which shows the linear relationship between the expected value in the control and the expected value in the other treatment.

Ignoring heterogeneity among the studies, this type of model is commensurate with a multiplicative model for m treatments of the form

$$\eta_{ij} = \alpha_j + \gamma_j \beta_i \quad (7a)$$

where α_j is an intercept for the j th treatment (analogous to $\theta + \tau_j$ of equation 4), and γ_j and β_i are fixed effects. As with equation 4, β_i is the effect of the i th study, and γ_j is the sensitivity of the j th treatment to the i th study (with $\gamma_j = 1$ indicating no sensitivity). One can also write equation 7a equivalently as

$$\eta_{ij} = \alpha_j + \beta_i + (\gamma_j - 1)\beta_i \quad (7b)$$

to show explicitly that the model encompasses main effects for treatment (α_j) and study (β_i), as well as an interaction effect $[(\gamma_j - 1)\beta_i]$. With just two treatments, rearranging and comparing coefficients gives the relationship between the parameters of equations 6 and 7:

$$\vartheta = \alpha_2 - \alpha_1(\gamma_2/\gamma_1) \quad \text{and} \quad \xi = (\gamma_2/\gamma_1) - 1$$

In general, for m treatments, one can write

$$\vartheta_j = \alpha_j - \alpha_1(\gamma_j/\gamma_1)$$

and

$$\xi_j = (\gamma_j/\gamma_1) - 1$$

One can choose any treatment as the control (treatment 1 is used here). Note that when the γ_j values are equal for the treatments, then $\xi = 0$ and the treatment effect is insensitive to the expected value of the control.

With among-study heterogeneity of treatment effects, the multiplicative model (equation 7a) is expanded to

$$\eta_{ij} = \alpha_j + \gamma_j \beta_i + u_{ij} \quad (8)$$

We call this model F3. Note that when $\gamma_j = 1$ for all treatments (i.e., for all j values), then model F3 reduces to model F1. Because of the multiplication of parameters, equation 7a or 8 is nonlinear. So, procedures for fitting nonlinear mixed models must be used for model fitting, such as adaptive Gaussian quadrature (Pinheiro and Bates 1995). The NLMIXED procedure in SAS can be used for fitting equation 8. Equation 7a is overparameterized, so constraints must be imposed on the parameters for model fitting. For our algorithm, we follow Ng and Grunwald (1997) and impose the identifiability constraints

$$\sum_{j=1}^m \gamma_j = m \quad \text{and} \quad \sum_{i=1}^K \beta_i = 0$$

Note that model fitting in NLMIXED is by ML (maximum likelihood) and not REML. The code is given in the Supplementary SAS program file. With a large number of studies (e.g., $K > 75$), and sparse study-by-treatment tables, model fitting may be quite challenging because of the large number of fixed-effect parameters (the number of β_i parameters increases with number of studies). Useful starting values for α_j are $\bar{\theta} + \bar{\tau}_j$ that are obtained from fitting the linear model F1; 1 is a useful starting value for all the γ_j terms.

TABLE 2

Restricted log-likelihood based measures of the fit of network meta-analytical mixed models for treatment means with either fixed (F1 to F3) or random (R1 to R6) study effects^a

Model	-2 Restricted log likelihood	AIC
F1	148.4 (-247.5) ^b	150.4 (36.5) ^b
F2	119.3 (-307.7)	129.3 (-17.7)
F3	- (-134.3)	(157.7)
R1	595.8	599.8
R2	585.7	599.7
R3	528.1	566.1
R4	588.2	602.2
R5	530.8	552.8
R6	560.8	582.8

^a Comparisons of restricted log-likelihoods or AIC statistics should only be among studies with a fixed study effect or a random effect.

^b -2 Log-likelihood and AIC based on maximum likelihood (ML) estimation are given in parentheses for the models with fixed main effect of study. Primary interest is in REML estimation, but comparisons of the fit of model F3 with models F1 and F2 must be done based on ML estimation. This is because model F3 was fitted using the NLMIXED procedure (which only uses ML). One cannot directly compare the statistics for ML and REML estimation.

One can also first fit equation 8 without the u_{ij} term, and use the estimated β_j s from this fit as the initial study-effect estimates when the full equation 8 is fitted to the data.

Example. With 136 studies, fitting model F3 to the wheat data were much more computationally intensive than fitting the other models. Although it only takes seconds to fit equation 8 to data sets with six treatments and under 30 studies on a 64-bit Windows 7 computer (*unpublished*), it took over 10 h to fit the same model to the wheat data with 136 studies. The estimated between-study variance was ~ 0 . The AIC for model F3 was considerably larger than found for model F1 (when F1 was refitted using ML instead of REML, so that direct comparisons can be made), suggesting that the multiplicative model did not provide a better fit than the simpler model for this example on wheat yield in relation to fungicide treatment. We found that the six estimated γ_j parameters only ranged from 0.97 to 1.03, indicating very little sensitivity of the treatment effects to the expected value in the control (results not shown). The estimates of $\xi_j = (\gamma_j/\gamma_1) - 1$, treating the control as treatment 1] were all about 0, further evidence for the lack of sensitivity.

Although a multiplicative model was not needed for this yield data set, Piepho et al. (2015) did find that the multiplicative model (equation 8) could provide a better fit of some medical data sets than an additive model. Fixed-effect multiplicative models—generally known as AMMI and GGE models—have also been found to be very useful for analyzing plant cultivar trials across multiple environments (locations, years) (Gauch 1988; Gauch et al. 2008; Malosetti et al. 2013; Yan et al. 2000). More details on the use of multiplicative models for meta-analysis can be found in Piepho et al. (2015). Several more meta-analytical data sets in plant pathology will need to be analyzed with this model to see how common it is for treatment effects to be proportional to the magnitude of the control mean.

LINEAR MODELS WITH RANDOM STUDY EFFECTS

General issues. It is possible to consider that study has a random effect on the summary result (mean) for each treatment and study, y_{ij} . The simplest situation would be for β_i in equation 4 to be a normally distributed random variable with mean of 0 and variance of $\sigma_{\beta_i}^2$; that is, $\beta_i \sim N(0, \sigma_{\beta_i}^2)$. This is analogous to the linear mixed model for incomplete block designs (Schabenberger and Pierce 2002) with a random block effect. With this model, information on treatment effects both within and between studies, that is, intrastudy and interstudy information, respectively, is recovered. This is generally considered highly desirable in single randomized experiments with blocking (Littell et al. 2006), and it can be justified based on the randomization of an incomplete block design. However, use of a random study effect in meta-analysis is controversial and has been criticized by some authors (Riley et al. 2008; Senn 2010; Senn et al. 2013) because there is typically no randomization among studies. For instance, the wheat plots in a study from a given state typically would not come from the same seed source as those used in a study in a different state; thus there would not be concurrent control. Nevertheless, several authors advocate use of random-study effects in meta-analysis (Arends et al. 2000; Greco et al. 2015; Stijnen et al. 2010; van Houwelingen et al. 2002). Based on the 650+ citations of the tutorial by van Houwelingen et al. (2002) in the Web of Knowledge, it is likely that many use this approach in practice. In fact, Senn (2010) believes that the interstudy information will be relatively low, so that one would not expect major differences in the fixed- and random-study situations. There have not been many comparisons of fixed versus random effects for study in multi-treatment meta-analysis in any area of science. Thus, we compared these here for the yield data set.

TABLE 3
Estimated contrasts of expected values ($\hat{\mu}_j - \hat{\mu}_1 = \hat{\tau}_j - \hat{\tau}_1$), with corresponding standard errors and significance levels (*P*) for selected network meta-analytical mixed models fitted to the treatment means for the wheat yield data set

Fungicide ^d	Model F1 ^a			Model F2 ^b			Model R1 ^c		
	Estimate	SE	<i>P</i>	Estimate	SE	<i>P</i>	Estimate	SE	<i>P</i>
Tebu. versus Control	0.274	0.026	<0.001	0.274	0.030	<0.001	0.274	0.026	<0.001
Prop. versus Control	0.198	0.036	<0.001	0.204	0.040	<0.001	0.199	0.036	<0.001
Proth. versus Control	0.427	0.037	<0.001	0.417	0.033	<0.001	0.429	0.037	<0.001
(Tebu.+Proth.) versus Control	0.432	0.031	<0.001	0.433	0.034	<0.001	0.433	0.031	<0.001
Met. versus Control	0.435	0.031	<0.001	0.441	0.031	<0.001	0.436	0.031	<0.001
(Tebu.+Proth.) versus Tebu.	0.157	0.031	<0.001	0.159	0.029	<0.001	0.158	0.031	<0.001
(Tebu.+Proth.) versus Met.	-0.003	0.034	0.925	-0.009	0.027	0.749	-0.003	0.034	0.921
	Model R2 ^e			Model R3 ^e			Model R5 ^e		
	Estimate	SE	<i>P</i>	Estimate	SE	<i>P</i>	Estimate	SE	<i>P</i>
Tebu. versus Control	0.269	0.026	<0.001	0.271	0.024	<0.001	0.273	0.024	<0.001
Prop. versus Control	0.194	0.036	<0.001	0.195	0.030	<0.001	0.198	0.031	<0.001
Proth. versus Control	0.426	0.037	<0.001	0.419	0.037	<0.001	0.425	0.040	<0.001
(Tebu.+Proth.) versus Control	0.429	0.031	<0.001	0.443	0.039	<0.001	0.443	0.039	<0.001
Met. versus Control	0.434	0.031	<0.001	0.450	0.033	<0.001	0.448	0.034	<0.001
(Tebu.+Proth.) versus Tebu.	0.160	0.030	<0.001	0.172	0.029	<0.001	0.170	0.030	<0.001
(Tebu.+Proth.) versus Met.	-0.004	0.033	0.893	-0.008	0.025	0.763	-0.005	0.027	0.867

^a $\hat{\sigma}_u^2 = 0.0253$.

^b $\hat{\sigma}_{u(1)}^2 = 0.0211$, $\hat{\sigma}_{u(2)}^2 = 0.0241$, $\hat{\sigma}_{u(3)}^2 = 0$, $\hat{\sigma}_{u(4)}^2 = 0.0183$, $\hat{\sigma}_{u(5)}^2 = 0.0016$, $\hat{\sigma}_{u(6)}^2 = 0.0620$.

^c $\hat{\sigma}_{\beta}^2 = 1.5382$, $\hat{\sigma}_u^2 = 0.02529$ (or $\hat{\sigma}_T^2 = \hat{\sigma}_{\beta}^2 + \hat{\sigma}_u^2 = 1.5635$; $\hat{\rho} = 1.5382/1.5635 = 0.984$).

^d Fungicide abbreviations: Tebu.: tebuconazole; Prop.: propiconazole; Proth.: prothioconazole; Met.: metconazole.

^e See Table 4 for structure of covariance matrix and Table 5 for the estimated variances and covariances in matrix format.

Use of random study effects may be reasonable when the studies are similar or preplanned, as was the case with our example. That is, a board of experts provided the basic design for individual studies involving the fungicide control of *Fusarium* head blight, and an effort was made nationally for collaborators to conduct the studies in a wide range of environments and cropping systems over several years. Moreover, it should be pointed out that analyses of crop varieties (genotypes) in multiple locations typically treat location (i.e., study) as a random effect (Denis et al. 1997). Although the studies are not randomly selected, one can assume that their effects are random (Higgins et al. 2009). Thus, we consider several versions of this form of model here, and compare the results with the fixed study situation.

Homogeneity of between-study variance. The basic version of equation 4 with random study effect ($\beta_i \sim N(0, \sigma_\beta^2)$) and random between-study treatment effect (i.e., treatment $\times$ study interaction) ($u_{ij} \sim N(0, \sigma_u^2)$) is referred to as model R1. A succinct alternative way of writing this and the subsequent models is obtained by defining the combined random effect of the study main effect and the study $\times$ treatment interaction, $h_{ij} = \beta_i + u_{ij}$ (Piepho 1999), giving the equation for the linear predictor as

$$\eta_{ij} = \theta + \tau_j + h_{ij} \quad (9)$$

The vector of the random effects for the i th study is given by $\mathbf{h}_i$. The variance-covariance matrix of $\mathbf{h}_i$ is specified by $\mathbf{G}$; that is, $\text{var}(\mathbf{h}_i) = \mathbf{G}$. With the assumption of normal distributions for the random effects, we can write $\mathbf{h}_i \sim N(\mathbf{0}, \mathbf{G})$. As with all the models in this paper, the summary result per study and treatment (y_{ij}) is given by equation 1 or 2, with the within-study variance-covariance matrix given by equation 3 (with $\mathbf{R}_i$ held constant in the analysis). This gives the following equation for y_{ij} :

$$y_{ij} = \theta + \tau_j + h_{ij} + e_{ij}$$

With expected values of the random effects equal to 0 (as defined above), the treatment mean is given by

$$\mu_j = \theta + \tau_j$$

The treatment effect relative to the control (if the control is the last treatment, and the parameterization of GLIMMIX is used) is simply τ_j . In general, as with the fixed-effects models, differences of expected values are given by $\mu_j - \mu_{j'} = \tau_j - \tau_{j'}$. The variance-covariance matrix of the estimated expected values, $\text{var}(\hat{\boldsymbol{\mu}})$, is a function of $\mathbf{R}_i$ and $\mathbf{G}$ (Littell et al. 2006).

All of the following models for a random study effect can be distinguished by the structure of $\mathbf{G}$. We consider six different structures in this paper. For model R1, $\mathbf{G}$ is shown in Table 4, where all matrix elements are determined by the block (σ_β^2) and between-study treatment effect (σ_u^2) variance components. Because all treatments within a study share the same random effect (β_i), all treatments are correlated, with a covariance for pairs of treatments given by the between-study variance (σ_β^2). With a total between-study variance of

$$\sigma_T^2 = \sigma_\beta^2 + \sigma_u^2$$

the between-study correlation for pairs of treatments is given by

$$\rho = \sigma_\beta^2 / \sigma_T^2$$

The structure of $\mathbf{G}$ is known as compound symmetry (CS). The $\mathbf{G}$ matrix could be equivalently written with σ_T^2 for the diagonal

TABLE 4
Between-study variance-covariance matrix for network meta-analytical mixed models of treatment means with random study effects

Model	$\mathbf{G} = \text{var}(\mathbf{h}_i)$	Parameters
R1	$\begin{pmatrix} \sigma_\beta^2 + \sigma_u^2 & \sigma_\beta^2 & \cdots & \sigma_\beta^2 \\ \sigma_\beta^2 & \sigma_\beta^2 + \sigma_u^2 & & \\ \vdots & & \ddots & \\ \sigma_\beta^2 & & & \sigma_\beta^2 + \sigma_u^2 \end{pmatrix}$	2
R2	$\begin{pmatrix} \sigma_{T(1)}^2 & \rho\sigma_{T(1)}\sigma_{T(2)} & \cdots & \rho\sigma_{T(1)}\sigma_{T(m)} \\ \rho\sigma_{T(2)}\sigma_{T(1)} & \sigma_{T(2)}^2 & & \\ \vdots & & \ddots & \\ \rho\sigma_{T(m)}\sigma_{T(1)} & & & \sigma_{T(m)}^2 \end{pmatrix}$	$m+1$
R3	$\begin{pmatrix} \sigma_{T(1)}^2 & \sigma_{T(12)} & \cdots & \sigma_{T(1m)} \\ \sigma_{T(21)} & \sigma_{T(2)}^2 & & \\ \vdots & & \ddots & \\ \sigma_{T(m1)} & & & \sigma_{T(m)}^2 \end{pmatrix}$	$m(m+1)/2$
R4	$\begin{pmatrix} \gamma_1^2 + \sigma_v^2 & \gamma_1\gamma_2 & \cdots & \gamma_1\gamma_m \\ \gamma_2\gamma_1 & \gamma_2^2 + \sigma_v^2 & & \\ \vdots & & \ddots & \\ \gamma_m\gamma_1 & & & \gamma_m^2 + \sigma_v^2 \end{pmatrix}$	$m+1$
R5	$\begin{pmatrix} \gamma_1^{(1)2} + \gamma_1^{(2)2} + \sigma_v^2 & \gamma_1^{(1)}\gamma_2^{(1)} + \gamma_1^{(2)}\gamma_2^{(2)} & \cdots & \gamma_1^{(1)}\gamma_m^{(1)} + \gamma_1^{(2)}\gamma_m^{(2)} \\ \gamma_2^{(1)}\gamma_1^{(1)} + \gamma_2^{(2)}\gamma_1^{(2)} & \gamma_2^{(1)2} + \gamma_2^{(2)2} + \sigma_v^2 & & \\ \vdots & & \ddots & \\ \gamma_m^{(1)}\gamma_1^{(1)} + \gamma_m^{(2)}\gamma_1^{(2)} & & & \gamma_m^{(1)2} + \gamma_m^{(2)2} + \sigma_v^2 \end{pmatrix}$	$2m$
R6	$\begin{pmatrix} \gamma_1^2 + \sigma_{v(1)}^2 & \gamma_1\gamma_2 & \cdots & \gamma_1\gamma_m \\ \gamma_2\gamma_1 & \gamma_2^2 + \sigma_{v(2)}^2 & & \\ \vdots & & \ddots & \\ \gamma_m\gamma_1 & & & \gamma_m^2 + \sigma_{v(m)}^2 \end{pmatrix}$	$2m$

elements and σ_{τ}^2 for the off-diagonal elements. This version is not used directly by GLIMMIX (or MIXED); this structure can be used directly with the HPMIXED procedure of SAS.

Heterogeneity of variances. As with fixed study effects, one can consider unequal between-study variances when there is an overall random study effect. There are two ways that model R1 could be generalized with an unequal diagonal elements of the **G** matrix, one is to have a constant between-study *correlation* for all pairs of treatments (ρ) with unequal between-study total variances ($\sigma_{T(j)}^2$), and the other is to have a constant between-study *covariance* (σ_{β}^2) and unequal σ_u^2 values ($\sigma_{u(j)}^2$). Both ρ and σ_{β}^2 cannot be simultaneously constant with unequal variances. Because of its standard usage in the GLIMMIX and MIXED procedures of SAS, we use the constant correlation version, and call this model R2. The **G** matrix is shown in Table 4, which is commonly called heterogeneous compound symmetry (CSH).

Unstructured variance-covariance matrix. Models R1 and R2 impose a high degree of structure on the variance-covariance matrix **G**. At the other extreme, one can allow the **G** matrix to be completely unstructured (UN), with a (potentially) different variance ($\sigma_{T(j)}^2$) for each treatment, and a (potentially) different covariance ($\sigma_{T(ij)}$) or correlation (ρ_{ij}) for each pair of treatments. We call the unstructured case model R3, which has the **G** matrix listed in Table 4.

Fitting a mixed model with an unstructured covariance matrix can be challenging, especially with a large number of treatments. Lack of convergence can be a problem, or the estimated **G** may not even be a valid variance-covariance matrix. To be a valid variance-covariance matrix, it must be at least positive semidefinite, a property in which all the eigenvalues of the matrix are nonnegative. Therefore, it is preferable to use an approximation (involving the same number of parameters) that guarantees that the estimated **G** is at least positive semidefinite. This can be done by reparameterizing **G** through its so-called Cholesky root; this approach is accomplished automatically using the GLIMMIX procedure with the type=chol option. Even with this approach, convergence can still be a problem, especially with large numbers of treatments (m), because the number of parameters increases as a quadratic function of m . The total number of parameters in **G** equals $m(m+1)/2$. When all treatments are in all studies, the number of studies must exceed the number of treatments to fit model R3 using likelihood methods; with a sparse data matrix (where only a subset of the treatments appear in each study), the number of studies needs to be considerably larger than the number of treatments in order to fit the model.

Model R3 was used by van Houwelingen et al. (2002), although they were dealing only with a small number of studies with two treatments. Analysis based on this model may be called multivariate meta-analysis because of the completely general correlation/covariance of the different treatment means. This model has found extensive use for meta-analysis in plant pathology, where the effects of different fungicides on crop disease control have been investigated (Ngugi et al. 2011b; Paul et al. 2008, 2010; Willyerd et al. 2012). Arends et al. (2003, page 1347) show how to use the elements of the **G** matrix to determine the relationship between the mean in the control and the treatment effect (say, $\tau_1 - \tau_m$); a plant pathology example is in Paul et al. (2008). This is one type of random-effects alternative to the approach taken with the fixed-effect multiplicative model (model F3).

Example. Code for fitting models R1 to R3 is given in Appendix A. Based on the AIC statistic, models R1 and R2 gave similar model fits, and model R3 with the unstructured variance-covariance matrix gave the best fit of these three (Table 2). A likelihood-ratio test for model R1 versus R3 also showed that R3 provided a better fit (difference = 67.7, df = 20, $P < 0.001$).

For model R1, the estimated total between-study variance for each treatment was $\hat{\sigma}_{\tau}^2 = 1.564$ and the estimated covariance for all

pairs of treatments was $\hat{\sigma}_{\beta}^2 = 1.538$. For model R2, the estimated total variance varied slightly with treatment, and consequently the estimated covariance depended (slightly) on the treatment pair (Table 5). The control treatment had the largest estimated variance (1.676). The estimated **G** matrix (Table 5) for model R3 also showed that the control (last treatment) had the highest between-study variance (1.673). The other variances and the covariances varied somewhat with treatment or treatment pair, but were fairly similar in value (Table 5). The covariances involving the control tended to be larger than the other covariances (see last row of estimated matrix). With the large covariances relative to the variances, the between-study correlations were all above 0.9.

Perhaps surprisingly for this example, the estimates of mean differences and the associated SEs were not greatly affected by the chosen model; results for models R1 to R3 are shown in Table 3. Consistent with the better fit of the unstructured model, there were greater differences in the magnitude of the SEs for model R3 than for R1. Results for pairwise mean differences were quite similar for the models with fixed study and random study effects; this can be seen by comparing models F1 and R1 in Table 3). Even though the estimated means had (considerably) larger SEs or variances for the models with random study effects compared with fixed study effects (*unpublished*), the SEs of the *differences* were quite similar. This is because the SE of a treatment effect is a function of the variances and covariance of the relevant treatment means. A large positive covariance can cancel out the large variances.

Multiplicative models for variability. The simple covariance structure of compound symmetry (models R1 or R2) may not adequately describe a data set, and an unstructured covariance matrix (model R3) may have an excessive number of parameters, based on the number of studies and treatments. Fortunately, there are several structures that are intermediate between what is found for models R1 (or R2) and R3. Multiplicative models for h_{ij} (Piepho 1997) offer a great deal of flexibility for representing covariance structure with relatively small number of parameters (especially compared with the unstructured case). These models have received considerable use in the analysis of genotype $\times$ environment data sets (Meyer 2009; Smith et al. 2005).

TABLE 5
Estimated variance-covariance matrix of h_{ij} (**G**; models R2, R3, and R5) for the wheat yield data set^a

Model	Variance-covariance estimates
R2	$\begin{pmatrix} 1.543 & & & & & \\ 1.516 & 1.536 & & & & \\ 1.519 & 1.516 & 1.543 & & & \\ 1.506 & 1.503 & 1.507 & 1.517 & & \\ 1.467 & 1.463 & 1.467 & 1.454 & 1.438 & \\ 1.583 & 1.579 & 1.583 & 1.570 & 1.528 & 1.676 \end{pmatrix}$
R3	$\begin{pmatrix} 1.526 & & & & & \\ 1.538 & 1.549 & & & & \\ 1.507 & 1.518 & 1.524 & & & \\ 1.507 & 1.519 & 1.521 & 1.528 & & \\ 1.471 & 1.482 & 1.484 & 1.479 & 1.479 & \\ 1.579 & 1.590 & 1.554 & 1.538 & 1.519 & 1.673 \end{pmatrix}$
R5	$\begin{pmatrix} 1.527 & & & & & \\ 1.523 & 1.535 & & & & \\ 1.508 & 1.512 & 1.525 & & & \\ 1.504 & 1.508 & 1.518 & 1.529 & & \\ 1.477 & 1.481 & 1.481 & 1.481 & 1.455 & \\ 1.581 & 1.585 & 1.549 & 1.539 & 1.522 & 1.676 \end{pmatrix}$

^a See Table 4 for the structure of the variance-covariance matrix for the different network meta-analytical mixed models with a random study effect.

We start by representing the treatment-effect variability term (h_{ij}) as a function of an underlying latent (i.e., unmeasured) variable for study: $h_{ij} = \gamma_j \beta_i + v_{ij}$. Now, β_i is a latent random variable for study effect (single value for all treatments in a study), γ_j is a parameter for the sensitivity of the j th treatment to the i th study, and v_{ij} is the unexplained variability term for the between-study treatment-effect (a random variable), with an assumption of $v_{ij} \sim N(0, \sigma_v^2)$. Because the $\gamma_j \beta_i$ multiplication term is overparameterized, a constraint must be placed on the parameters. For identifiability purposes, the variance of β_i is constrained to 1 (that is, $\beta_i \sim N(0, 1)$). The model for the linear predictor can be written as

$$\eta_{ij} = \theta + \tau_j + \gamma_j \beta_i + v_{ij} \quad (10)$$

We call this model R4. The $\mathbf{G}$ matrix for the variance-covariance of $\mathbf{h}_i$ is shown in Table 4.

Model R4 (equation 10) is the random-study version of the multiplicative model F3. That is, $\theta + \tau_j$ and α_j both represent the expected linear predictor for the j th treatment, and the $\gamma_j \beta_i$ term represents the interaction of study and treatment as a multiplicative term. Unlike in model F3, the study effect here is a random variable; for both models, γ_j represents the sensitivity of the j th treatment to the i th study. Readers should see Piepho et al. (2015) for applications of this model in meta-analysis.

Model R4 has the same number of variance-covariance parameters in the $\mathbf{G}$ matrix as found for model R2 with the CSH structure ($m + 1$, corresponding to the m γ_j values [one for each treatment], plus σ_v^2) (Table 4). However, the structure of $\mathbf{G}$ matrix is different for the two models. For instance, the intraclass correlation is not held constant with model R4

$$\rho_{12} = (\gamma_1 \gamma_2) / ((\gamma_1^2 + \sigma_v^2)^{1/2} (\gamma_2^2 + \sigma_v^2)^{1/2})$$

whereas it is constant by definition with model R2. All between-study variances, covariances, and correlations are easily obtained from the GLIMMIX output when a multiplicative parameterization (such as in model R4) is used. See Appendix A for details.

Expanded multiplicative models for variability. Model R4 can be expanded in various ways to accommodate more complex structures for $\mathbf{G}$. These expansions can be seen as attempts to adequately describe a complex variance-covariance structure without using the large number of parameters in model R3. One generalization is obtained by expanding the model for h_{ij} to have two multiplication terms:

$$h_{ij} = \gamma_j^{(1)} \beta_i + \gamma_j^{(2)} \beta_i + v_{ij}$$

where $\gamma_j^{(1)}$ and $\gamma_j^{(2)}$ are the sensitivities to the studies (i.e., the sensitivities of treatment j to study i). Essentially, some of the unexplained variability of model R4 (σ_v^2) is (potentially) accounted for by a second multiplicative term for the sensitivity of treatment to the study. This gives the model for the linear predictor:

$$\eta_{ij} = \theta + \tau_j + \gamma_j^{(1)} \beta_i + \gamma_j^{(2)} \beta_i + v_{ij} \quad (11)$$

We call this model R5. The distribution assumptions are the same as for model R4 ($\beta_i \sim N(0, 1)$, $v_{ij} \sim N(0, \sigma_v^2)$). The $\mathbf{G}$ matrix corresponding to equation 11 is given in Table 4. Additional constraints are required for identifiability (i.e., for parameter estimation). In the GLIMMIX procedure of SAS, the γ parameter for the first treatment ($j = 1$) of the second dimension is set to 0 (i.e., $\gamma_1^{(2)} = 0$). With this constraint, there are $2m$ parameters in $\mathbf{G}$.

Even though models R4 and R5 allow for unequal treatment variances if the γ_j values are not the same (see the diagonal terms in the $\mathbf{G}$ matrix in Table 4), one can accommodate additional

heterogeneity by relaxing the assumption about the variance of v_{ij} . That is, one can start with the $\mathbf{G}$ structure of model R4 (one multiplicative term) but allow σ_v^2 to take a (potentially) different value for each treatment ($\sigma_{v(j)}^2$). We call this model R6, with the $\mathbf{G}$ matrix given in Table 4. Accommodating the variance heterogeneity increases the number of parameters in $\mathbf{G}$ to $2m$, the same as in model R5, but with a different structure.

Example. Models R4 to R6 were fitted using the code described in Appendix A, with additional code in the supplementary SAS program file. Model R5, with two multiplicative terms, had the lowest AIC of all the random-study models tested with the example data set (Table 2). The estimates of the $\gamma_j^{(1)}$ parameters were all about the same (a result similar to that found for model R4 [L. V. Madden, unpublished data]); the estimates of the $\gamma_j^{(2)}$ parameters were erratic, and could not be easily interpreted. With the constraint on the $\gamma_j^{(2)}$ estimates for this model, sensitivity parameters beyond the first dimension are difficult to interpret; one can simply think of the $\mathbf{G}$ structure for model R5 as a convenient low-rank approximation to the unstructured variance-covariance matrix of model R3 (compare the estimated $\mathbf{G}$ matrices for models R3 and R5 in Table 5). The largest between-study variance for both models was for the control (last treatment) and the smallest variance was for the next to last treatment. Due to the larger variance for the control, the covariances generally were largest for the control and the other treatments (last row of the matrix). However, the ranges in variances or covariances were not large. With other data sets, we found a much greater range in the between-study variances and covariances (L. V. Madden, unpublished data).

The estimated treatment contrasts and standard errors for model R5 are shown in Table 3. Clearly, in this example, this random-effect structure did not have a large effect on the results for these selected contrasts. In some cases, the estimated standard error was slightly larger than for model R1 (or R3), and in other cases the estimates were lower. Conclusions were not affected here, although with other data sets, this would not necessarily be so. Madden et al. (2014) found model R5 to be very useful in a meta-analysis of over 300 studies with 27 treatments for the control of Fusarium head blight of wheat. In fact, it was not possible to fit the unstructured model R3 with the large number of treatments to this larger data set, even though there were considerably more studies than treatments.

MODERATOR VARIABLES

The among-study variance, σ_u^2 , or the elements of the $\mathbf{G}$ variance-covariance matrix (Table 4), reflect the heterogeneity of treatment effects on the estimated effect size across studies. There can be multiple causes of this heterogeneity, such as the diversity in the ways the studies were conducted and other characteristics of the studies (Borenstein et al. 2009; Lipsey and Wilson 2001; Madden and Paul 2011; Paul et al. 2007, 2011). For the example data set in this paper, the environment clearly varied; although susceptible cultivars were always grown, the specific cultivars grown depended on the location. One may be able to account for study-specific effects by incorporating so-called moderator variables in the meta-analytical model (van Houwelingen et al. 2002; Whitehead 2002). A moderator variable is a characteristic of each study that can be used in analysis. These variables can be categorical (e.g., resistance level) or continuous (e.g., amount of rainfall during a critical time of the growing season). We demonstrate the analytical approach with a categorical variable.

For the wheat yield data set, studies were either conducted with spring or winter wheat market-class cultivars, depending on the location (Paul et al. 2008, 2010). We expand on model R1 here to incorporate the moderator-variable (wheat market class) effect. The other fixed or random study effect models can all be expanded in the

same way. We need to add an additional subscript index to y (l) or to some of the other terms in the model in order to represent the specific level of the moderator variable; we define y_{ijl} as the mean for the i th study, j th treatment, and l th level of the moderator variable (in this case, $l = 1$ for spring wheat and $l = 2$ for winter wheat). One version of the expanded model can be written as

$$y_{ijl} = \theta + \tau_j + h_{ij} + \delta_{jl} + e_{ijl} \quad (12a)$$

where δ_{jl} is the effect of the l th wheat class ($l = 1, 2$) and j th treatment on the effect size (i.e., mean). We can also write $y_{ijl} = \eta_{ijl} + e_{ijl}$ (see equation 1 for the situation without a moderator variable) and define the linear predictor as

$$\eta_{ijl} = \theta + \tau_j + h_{ij} + \delta_{jl} \quad (12b)$$

Because no main effect of wheat class is included, δ_{jl} is simultaneously indicating the effect of wheat class on the study mean and the effect of wheat class on the treatment effect. Alternatively, main effect and interaction terms could be used:

$$y_{ijl} = \theta + \tau_j + h_{ij} + \chi_l + \delta_{jl} + e_{ijl} \quad (12c)$$

where χ_l is the main effect of wheat class (indicating overall effect of l th wheat class on the effect size) and δ_{jl} is the interaction (now specifically indicating the effect of wheat class on the effect of treatment). We prefer the simpler equation 12a. Significance of the moderator is determined with a type 3 test of δ (Littell et al. 2006). If study is a fixed effect (such as in model F1), then the expanded model is:

$$y_{ijl} = \theta + \beta_{il} + \tau_j + u_{ij} + \delta_{jl} + e_{ijl} \quad (13)$$

where β_{il} is effect of study nested within wheat market class. Appendix A presents the GLIMMIX code to fit models R1 and F1 with a moderator variable. The supplementary SAS program file demonstrates how to obtain the expected values of differences of treatment means for each level of the moderator variable; e.g., $\mu_{ji} - \mu_{mi}$ for the difference between the j th and the m th treatment means at level l of the moderator variable.

As shown in Paul et al. (2010), using model R3, wheat market class had a strong effect on the treatment effects, with generally larger effects of treatments (relative to the control) with spring wheat compared with winter wheat. Similar results were found with the other models discussed here (*unpublished*). For instance, when fitting equation 12a to the example data set for yield, the F test for the moderator effect (δ) was 4.19 ($P < 0.001$). When fitting model R5 with a moderator variable term, the F test for the moderator effect was 3.89 ($P < 0.001$).

INCONSISTENCY

As discussed at the start of the paper, network meta-analysis often provides direct and indirect evidence on the contrasts of interests, such as the pairwise difference of two means. This is a consequence of the variability in the actual treatments that comprise individual studies. When direct comparisons of means are different from indirect comparisons, then the network is inconsistent, based on the concepts in Lu and Ades (2004). There has been considerable effort made in medical statistics to quantify inconsistency in multi-treatment meta-analysis based on the discrepancy of direct and indirect evidence (Dias et al. 2013a; Donegan et al. 2013; Higgins et al. 2012; Krahn et al. 2013; Lu and Ades 2004; Salanti et al. 2008; Tu 2014). This analysis is quite complex and requires great care in tracing the sources of inconsistency through loops in the network and specifying the required dummy variables and covariance matrix in the model (Piepho 2014; Piepho et al. 2012b).

Piepho et al. (2012b) extended the concept of inconsistency and developed a modeling approach that is much easier to implement with standard mixed modeling software such as the GLIMMIX or MIXED procedures of SAS. Related approaches have been introduced by Greco et al. (2015). An inconsistency of direct and indirect comparisons implies that treatment differences depend on the design of the study (where design indicates here the set of treatments in the study). Thus, an appropriately defined interaction term can be used to represent the inconsistency if it exists (Piepho 2014).

We first categorize all studies based on the treatments in the study, and identify the design of the study with the k subscript ($k = 1, \dots, D$), where D is the number of unique designs. We then fit the following model (starting with model F1):

$$\eta_{ijk} = \theta + \omega_k + \beta_{ik} + \tau_j + (\tau\omega)_{jk} + u_{ijk} \quad (14)$$

where ω_k is the main effect of the k th design (a fixed effect), β_{ik} is the effect of the i th study nested in the k th design (a fixed effect, equivalent to an interaction of study and design), $(\tau\omega)_{jk}$ is the interaction of the j th treatment and the k th design (fixed effect), and u_{ijk} is the random term for the variability in treatment effect between studies nested within design (equivalent to the interaction of design, study, and treatment). A significant $\tau\omega$ interaction (say, from a Wald test) is an indication of inconsistency. This form of inconsistency includes any effect of study design on the treatment effects, and is not restricted to inconsistency that is exhibited when direct comparisons are different from indirect comparisons.

Piepho (2014) gives a detailed tutorial on the characterization of inconsistency in network meta-analysis using the mixed-model framework of equation 14, and also shows how one can determine the contribution of each design to the overall inconsistency of the network. Piepho et al. (2015) further extend equation 14 by use of a multiplicative term for the interaction. To date, appraisal of inconsistency has not been a major issue in plant pathology, but we expect this will become more common as meta-analysis continue to grow in popularity.

Example. There were 13 unique designs for the yield data set (see labels in SAS program file). For instance, there were 23 studies that included all six treatments (including the control). We call this design 1. Twenty-one studies had only the control and tebuconazole treatment; we call this design 4. Upon fitting equation 14, the F test for the design $\times$ treatment interaction ($\tau\omega$) was only 0.82 ($P = 0.756$); thus, there was little evidence for inconsistency with this network. The F test was about the same when a random study effect model was used. Code is given in the SAS program file.

DISCUSSION

Network (or multi-treatment) meta-analysis “incorporates all available evidence into a general statistical framework for comparisons of all available treatments to resolve the limitations of traditional pairwise meta-analyses” (Tu 2014). Not surprisingly, the numbers of papers using network meta-analysis have grown substantially over the past decade (Cipriani et al. 2013; Lee 2014), especially in medicine. The increased popularity of network meta-analysis is due, at least in part, to the influential papers by Lu and Ades (2004, 2006). Although these authors took a Bayesian approach, others have shown how to use frequentist methods for this type of analysis (Greco et al. 2015; Jones et al. 2011; Krahn et al. 2013; Lu et al. 2011; Piepho et al. 2012b, 2015; Senn et al. 2013; Tu 2014). An extremely nice feature of the multi-treatment approach is that all studies do not have to contain all treatments (or effect sizes, in general). A challenge with the multi-treatment approach is that the analytical methods are more complex, leading to confusion about how to choose models, fit them to the data, and interpret the results.

In this review, we presented a collection of mixed models that are useful for network meta-analysis in plant pathology or other disciplines; Appendix A and the online supplemental files provide the necessary code to fit these models in SAS using likelihood methods. We show the models for the analysis of summary results (e.g., estimated means) for each study rather than for individual-plot (replicate) data analysis, because the original observations are often not available for all studies. Fortunately, results based on individual-plot data and summary treatment means per study are typically quite similar (Möhring and Piepho 2009; Piepho et al. 2012b). It is not common to compare the fit of different models for network meta-analysis, so we fitted all of the presented models to an example published data set on the effects of fungicide treatment on the control of Fusarium head blight, with yield as the response variable (Paul et al. 2010). Meta-analysis has been frequently used to synthesize the results from Fusarium head blight studies (Madden and Paul 2011; Madden et al. 2014; Paul et al. 2006, 2007, 2008). Past research showed that prothioconazole, tebuconazole + prothioconazole, and metconazole provided the greatest degree of disease and toxin reduction and the highest yield increases of the treatments considered. The treatment effects on yield were confirmed here using several different models, most of which were not used previously in the published papers on this disease. Use of a model with an unstructured between-study covariance matrix (model R3), which accounts for a heterogeneity in variances and covariances of the treatment effects, is popular for meta-analysis in plant pathology (Ngugi et al. 2011b; Paul et al. 2008, 2010; Willyerd et al. 2012). However, this matrix is quite complex, with the largest number of parameters of the covariance matrices being considered; model fitting can be difficult, even with algorithms that ensure that the matrix is at least positive semidefinite. With the example data set, a less complex covariance matrix (model R5), based on a multiplicative structure, provided the best fit; this model still accounted for the unequal between-study variances and unequal covariances (Table 4). Interestingly, several models, with either a fixed or random study effect, gave similar results for treatment effects for this example. Apparently for this data set on wheat yield, the differences in between-study variances (and covariances) among the treatments were not sufficient to greatly affect the estimated differences in treatment means and their associated standard errors. Sensitivity of the results to the selected model will depend on the features of the data set. For instance, when the response variable was a transformation of Fusarium head blight field severity ("disease index") or deoxynivalenol concentration in the harvested grain (Paul et al. 2008), some of the estimated treatment effects depended on the model being fitted (L. V. Madden, *unpublished data*). It is recommended that more such comparisons be done for future meta-analyses rather than simply selecting one model a priori.

There are several issues that are open to judgment when analyzing multi-treatment data. For instance, should the study effect be fixed or random in the model? Senn (2010) and Senn et al. (2013) certainly make a compelling case based on statistical considerations that study should be a fixed effect. Nevertheless, the approach advocated by van Houwelingen et al. (2002), which considers study a random effect, is very popular in meta-analysis. The random-study case is generally followed in plant pathology. Senn (2010) even acknowledges that the differences in the results between random and fixed study effects will not be substantial, in general. This is because a random study effect behaves like a fixed study effect when the random effect has a large variance (relative to the residual variances) (Searle et al. 1992). We found the similarity in results between the fixed and random study effect to be the case for the example, and for other data sets (L. V. Madden, *unpublished data*). We have also noticed that it is easier to fit models (i.e., achieve convergence in the iterative restricted likelihood algorithm) with a random study effect to large data sets (*unpublished*).

Either the conditional model of Lu and Ades (2004, 2006), defined in terms of treatment contrasts with a baseline, or the unconditional mixed model, defined in terms of treatment means, can be used for network meta-analysis. In practice, the two modeling approaches can be equivalent (Piepho 2014; Piepho et al. 2012b). In particular, the unconditional model with a fixed study effect (F1) presented here is equivalent to the commonly used conditional model of Lu and Ades (Piepho 2014; Piepho et al. 2012b). When analyses are done on treatment means per study (i.e., a two-stage analysis), the results are identical. We showed this with the example in Appendix B. It should be emphasized that it is much easier to fit the unconditional models to data (Greco et al. 2015; Piepho et al. 2012b). Adding moderator-variable effects to unconditional models is also straightforward (Paul et al. 2010), as is adding terms to quantify inconsistency in the network (Piepho 2014). Jones et al. (2011) and Tu (2014, 2015) give other elaborations of the Lu and Ades models, mostly for individual-participant-(replicate)-data (IPD) situations. In the models of Jones et al. (2011), additional constraints are placed on the u_{ij} random effects (equation 4) so that they sum to 0. Tu proposes to center the dummy variables within each study for treatment effects in the model. However, the equivalence of the conditional and unconditional models does not depend on these constraints. More details on analytical approaches to multi-treatment problems can be found in the special issue of *Research Synthesis Methods* (Salanti and Schmid 2012), with an emphasis on the conditional modeling framework. No matter what modeling approach is used, we feel that more effort should be made in plant pathology to determine whether or not treatment effects are consistent in the network (Lu and Ades 2006; Piepho 2014; Piepho et al. 2012b, 2015). We demonstrated the inconsistency-testing approach of Piepho et al. (2012b), which is easy to implement for unconditional models using mixed-modeling software.

Clearly, based on recent trends in other disciplines (Lee 2014), plant pathologists and other plant protection scientists will be increasingly using network meta-analysis in the coming years. It is hoped that this review provides a framework for reasonable modeling of the data and interpretation of the results.

APPENDIX A: SOFTWARE CODE

Model F1 (fixed study effects, homogeneity). Equation 5 can be fitted using restricted maximum likelihood (REML) or maximum likelihood (ML) with the MIXED or GLIMMIX procedure of SAS, and with other programs. An advantage of GLIMMIX is that there are options for different optimization methods, of great use for difficult-to-fit problems. Generic code is given here for GLIMMIX. The supplementary SAS files provide specific code for the examples, and additional code for obtaining contrasts.

There needs to be a separate record for each treatment (treat) of each study (study), with variables for y_{ij} (response) and s_{ij}^2 (sampvar). The tricky issue is fixing the within-study variances in the model fitting process. There are different ways of doing this, which include writing out each sampling variance in a PARMS statement, or placing the within-study variances in a separate file, and using the pdata=option to identify the file (van Houwelingen et al. 2002). However, there is a generally much easier way of doing this (Piepho 1999; see also online supplementary files of Madden and Paul 2011). One defines the observation weights as $1/s_{ij}^2$ (wt=1/sampvar), and determines wt in a data step. One then holds the residual variance (the last variance term in the model when there is more than one variance) at the value of 1 using options on the parms statement. The combined effect of using the weight and parms statements (with the hold= [or eqcons=] option) is to fix the diagonal elements of the $\mathbf{R}_i$ variance-covariance matrix (equation 3) at the $1/\text{wt}$ (i.e., s_{ij}^2) values (see page 159 of Piepho 1999). The output indicates a residual of 1, but it is actually $1/\text{wt}$ (different for each study).


```
proc glimmix data=a method=rspl;
class study treat;
weight wt;
model response = study treat;
random treat / subject=study;
parms (1) (1) / hold=2;
lsmeans treat / diff cl;
run;
```

The method=rspl option specifies REML estimation in this procedure. The “1” for the first term in the parms statement is the starting value for the variance of u_{ij} in the iterative model-fitting algorithm. The second “1” in the parms statement is the “variance” of e_{ij} (held fixed with the hold=2 option). It is critical that the last variance term in the parms statement be held fixed at 1 (coupled with a weight statement) for this and subsequent models. Investigators who use different software for fitting model F1 (and the subsequent models) must ensure that they are using code to hold the within-study variances at fixed values, otherwise the model fitting will be in error.

Model F2 (fixed study effects, heterogeneity). For the situation with six treatments, one can fit model F2 with GLIMMIX by changing the random and parms statements to

```
random treat / subject=study group=treat;
parms (1) (1) (1) (1) (1) (1) (1) / hold=7;
```

The group=option on the random statement specifies a separate between-study variance for each treatment. With six treatments, there are six between-study variances and the residual, the latter being held at 1. Thus, the parms statement has seven total variance terms, and the seventh (last) one is fixed at 1 (using hold=7). This approach of fixing the residual at 1 must be coupled with the use of the weight statement. One may have to explore different starting values for the nonresidual variance parameters (all the variance terms except the last one in the parms statement) in order to achieve convergence. Other statements would be unchanged.

Model F3. This model requires extensive coding in the NLMIXED procedure to fit to data. Refer to the provided SAS file for details.

Model R1 (random study effects, homogeneity). With the GLIMMIX procedure of SAS, one can use a modification of the program given for model F1 to fit the model with a random effect for study:

```
proc glimmix data=a method=REML;
class study treat;
weight wt;
model response = treat;
random treat / subject=study type=CS g gcorr;
parms (1) (1) (1) / hold=3;
lsmeans treat / diff cl;
run;
```

The random statement is specifying the variance-covariance matrix of the $\mathbf{h}_i$ vector (sum of study main effect and interaction of study and treatment) ($\mathbf{G}$). In this compound symmetry (CS) syntax of the procedure, the σ_{β}^2 term (Table 4) is considered as a covariance, and thus it can take negative values (although usually the estimate is nonnegative). In the output, σ_{β}^2 is listed as “CS” and σ_u^2 is listed as “Variance” (with study for the subject). The g and gcorr options specify that the estimated $\mathbf{G}$ variance-covariance matrix, as well as the correlation matrix derived from the estimated $\mathbf{G}$, be displayed. Note that there are now three random-effect parameters, including

the within-study variance (which is the third parameter). The within-study variance is being held fix using hold=3 option (this is coupled with the use of the weight statement as described above for model F1).

For model R1, one could alternatively write the random statement as

```
random intercept treat
/ subject=study type=VC;
```

for a so-called variance component (VC) version of the CS model. Here, σ_{β}^2 is a variance (listed as Intercept, with study as the subject), so that only nonnegative values are definable (with the default syntax). The two versions will give the same model fit unless the covariance in the first formulation is negative. Note that the ρ estimate must be calculated by hand for this VC structure in GLIMMIX based on the estimates of the two variances. For the VC formulation, one can also allow for negative “variances” by using the NOBOUND option in the GLIMMIX statement; in this case, the two formulations are equivalent in all cases.

Model R2 (random study effects, heterogeneity). Fitting the heterogeneous compound symmetry (CSH) model R2 is straightforward by modifying the random and parms statements in GLIMMIX. With six treatments, there are eight variance-covariance parameters (six treatment variances, a correlation parameter, and the within-study variance) (Table 4). This formulation assumes, as used above, that a weight statement is used that defines the inverse of the within-study variance, and that the residual variance is held constant at 1.

```
random treat / subject=study
type=CSH g gcorr;
parms (1) (1) (1) (1) (1) (1) (.5) (1) / hold=8;
```

Note that the correlation term is next to last, and its initial estimate must be less than 1. Unlike in the case with the simpler CS structure (model R1), the GLIMMIX procedure directly estimates ρ (labeled “CSH” in the output), as well as the treatment-specific total variances (labeled “Var(1)”, etc., for $\sigma_{T(1)}^2$, etc.).

Model R3 (random study effects, unstructured variance-covariance matrix). Fitting model R3 is done by modifying the random and parms statements in GLIMMIX. With six treatments, there are 22 variance-covariance parameters (six treatment variances, 15 covariance parameters, and the within-study variance (parameter 22) that is held constant at 1. As always, fixing the within-study variance at 1 must be coupled with a weight variable that is the inverse of the within-study variances in order to obtain a fixed within-study variance that varies with study. One can use type=UN (or type=UNR) on the random statement to specify an unstructured $\mathbf{G}$, but as discussed in the paper, a better approach is to use CHOL to constrain $\mathbf{G}$ to be at least positive semidefinite through a Cholesky formulation.

```
random treat / subject=study
type=CHOL g gcorr;
parms (1) (1) (1) (1) (1) (1) (1) (1) (1) (1)
(1) (1) (1) (1) (1) (1) (1) (1) (1) (1)
(1) (1) / hold=22;
*-Last term is held at 1;
```

In the MIXED procedure, one uses type=FA0(6) to obtain the same matrix, where the “6” corresponds to the six treatments.

Model R4 (random study effects, first-order multiplicative structure). One fits model R4 to the data by using GLIMMIX with the following code for the random and parms statements (assuming that the weight statement is used with the inverse of the sampling variances for each study).


```
random treat / subject=study
type=FA1(1) g gcorr;
parms (1) (1) (1) (1) (1) (1) (1)
(1)/hold=8;
```

Note that the random multiplicative structure is considered a 1st order factor analytical structure [FA1(1)]. For six treatments, the parms statement refers to the seven parameters of the **G** matrix and the residual variance held constant.

Model R5 (random study effects, second-order multiplicative structure). For the model with two multiplicative terms, one uses the following code in the GLIMMIX procedure, using the FA1(2) covariance structure.

```
random treat / subject=study
type=FA1(2) g gcorr;
parms (1) (1) (1) (1) (1) (1) (1) (1)
(1) (1) (1) (1) (1)/hold=13;
```

For six treatments, the parms statement refers to the six $\gamma_j^{(1)}$ parameters, the five $\gamma_j^{(2)}$ parameters (as described above, $\gamma_1^{(2)}$ is constrained to 0), and the σ_v^2 variance of the **G** matrix, plus the residual variance held constant.

Model R6 (random study effects, first-order multiplicative structure, heterogeneity). The multiplicative model with unequal variances is fitted to the data by changing FA1(1) to FA(1). The random and parms statements (assuming that the weight statement is used with the inverse of the sampling variances for each study) are now

```
random treat / subject=study
type=FA(1) g gcorr;
parms (1) (1) (1) (1) (1) (1) (1) (1)
(1) (1) (1) (1) (1)/hold=13;
```

For six treatments, the parms statement refers to the six γ_j parameters and six variances ($\sigma_{\gamma_j}^2$) of the **G** matrix, plus the residual variance held constant. It may be necessary to try different starting values for the random-effect parameters (except for the residual [last] term) in order to get convergence.

Additional models. For random study models with categorical moderator variables, one modifies the GLIMMIX statements to

```
class study treat moder;
model response = treat treat*moder / ddfm=KR;

When study is a fixed effect, one modifies the model statement to

model response = study(moder) treat
treat*moder / ddfm=KR;
```

because studies are nested within levels of the moderator. When fixed effects (such as moderator variables) are added to the model, one should override the default denominator degrees of freedom method and use the Kenward-Roger (KR) method. This is accomplished by adding the ddfm=KR option to the model statement. One uses contrast, estimate, or lsmeans statements to obtain differences of treatment means for different levels of the moderator. An example is in the SAS program file.

For a test of inconsistency, one first defines the categorical variable group to indicate the design of the study (indicator of treatments present in the study). For study as a fixed effect (model F1), one modifies the GLIMMIX code to

```
class study treat group;
model response = group study(group)
treat treat*group / ddfm=KR;
```

The test of the interaction of treatment and group (treat*group) is used for determining inconsistency. For models with moderator variables or with the design variable, one must still use the weight statement and hold the last variance at a fixed value of 1 (as described above). One needs to use the same random statements as described for previous models.

For models based on within-study contrasts (conditional models as described in Piepho et al. 2012b), one should refer to the contrast SAS program file.

APPENDIX B: CONTRAST-BASED MODELING

Piepho et al. (2012b) and Piepho (2014) show that model F1 gives identical results to the contrast-based (conditional) model when summary results are analyzed in the second stage. We demonstrate this here for the situation where the same baseline (control) is present in each study. Methodology is considerably more tedious when the baseline varies among studies.

Assuming that the control is the last (m th) treatment, one can write the estimated effect size for the i th study as the pairwise treatment difference: $z_{ij} = y_{ij} - y_{im}$, for $j = 1, \dots, m-1$. The response variable in the analysis is now the contrast with the control. Substituting equation 1 for y_{ij} and y_{im} into z_{ij} (with η_{ij} given by equation 4), one obtains

$$z_{ij} = (\tau_j - \tau_m) + (u_{ij} - u_{im}) + (\theta - \theta) + (\beta_i - \beta_i) + (e_{ij} - e_{im}) \\ = \tilde{\tau}_j + \tilde{u}_{ij} + \tilde{e}_{ij} \quad (\text{B1})$$

as a model for the treatment contrasts. The main effect of study in equation 4 (β_i) cancels out. Here, $\tilde{\tau}_j (= \tau_j - \tau_m)$ represents the treatment effects (expected differences between each treatment and the baseline). $\tilde{u}_{ij}$ ($u_{ij} - u_{im}$) is the between-study random effect for the contrasts, and $\tilde{e}_{ij}$ ($e_{ij} - e_{im}$) is the within-study random effect for the contrasts (the latter being held fixed in the analysis). Defining $\tilde{\mathbf{u}}_i$ and $\tilde{\mathbf{e}}_i$ as the vectors of the random effects in the i th study, we can succinctly represent the distributional assumptions as: $\tilde{\mathbf{u}}_i \sim N(0, \tilde{\Sigma})$ and $\tilde{\mathbf{e}}_i \sim N(0, \tilde{\mathbf{R}})$. We call this model B1.

The expected value of the treatment effect (relative to the control) is given by $\tilde{\mu}_j = \tilde{\tau}_j$. Other treatment effects, (say, j versus j') can be determined by subtracting the relevant $\tilde{\mu}_j$ values. Assuming that treatments are independent within studies (i.e., starting with the distributional assumptions of model F1), the within-study covariance matrix for the $m-1$ contrasts is given by

$$\tilde{\mathbf{R}}_i = \begin{pmatrix} s_{i1}^2 + s_{im}^2 & s_{im}^2 & \cdots & s_{im}^2 \\ s_{im}^2 & s_{i2}^2 + s_{im}^2 & & \vdots \\ \vdots & & \ddots & \\ s_{im}^2 & & & s_{im-1}^2 + s_{im}^2 \end{pmatrix} \quad (\text{B2})$$

One should compare this with the diagonal matrix in equation 3. The within-study sampling variance of each contrast is the sum of the variances for each treatment; the covariance for all contrasts is now nonzero and equal to the variance for the baseline treatment (which is common to all contrasts). The between-study covariance matrix is given by

$$\tilde{\Sigma} = \begin{pmatrix} 2\sigma_u^2 & \sigma_u^2 & \cdots & \sigma_u^2 \\ \sigma_u^2 & 2\sigma_u^2 & & \\ \vdots & & \ddots & \\ \sigma_u^2 & & & 2\sigma_u^2 \end{pmatrix} = \sigma_u^2 \begin{pmatrix} 2 & 1 & \cdots & 1 \\ 1 & 2 & & \\ \vdots & & \ddots & \\ 1 & & & 2 \end{pmatrix} \quad (\text{B3})$$

This can be compared with the Σ matrix in Table 1 for model F1. Now, the between-study variance for z_{ij} is double the value for y_{ij} ,

and there is a nonzero between-study covariance for all contrasts equal to half of the among-study variance for z_{ij} .

Although one can write the matrix equations very easily, fitting model B1 requires much more effort than fitting model F1 or any of the other models discussed in this paper. One has to construct the two covariance matrices for each study and place these in data files, and then link these to the summary-result data set when using GLIMMIX in SAS. This is complicated by the fact that the studies often have different treatments. When the baseline is not always the same, one also has to use a more general form of model B1 and define dummy variables for the contrasts within studies (Piepho et al. 2012b).

An argument in favor of model B1 (or similar models based on contrasts within studies) is that the analysis respects the principle of concurrent control (Senn 2010), so interstudy information is not influencing the estimated treatment effects (Ohlssen et al. 2014). However, the same thing can be accomplished by treating study as a fixed effect and directly analyzing treatment summaries for each study (such as with models F1 and F2). We show the contrast-based methods for our example in an online supplementary SAS file. As required, the results (estimated treatment effects, standard errors, among-study variance) are identical to fitting model F1 (equation 4) to the data, and are not repeated here because of space considerations.

LITERATURE CITED

- Arends, L. R., Hoes, A. W., Lubsen, J., Grobbee, D. E., and Stijnen, T. 2000. Baseline risk as predictor of treatment benefit: three clinical meta-re-analyses. *Stat. Med.* 19:3497-3518.
- Arends, L. R., Vokó, Z., and Stijnen, T. 2003. Combining multiple outcome measures in a meta-analysis application. *Stat. Med.* 22:1335-1353.
- Bagos, P. G. 2008. A unification of multivariate methods for meta-analysis of genetic association studies. *Stat. Appl. Genet. Mol. Biol.* 7:31.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., and Rothstein, H. R. 2009. *Introduction to Meta-Analysis*. John Wiley & Sons, Chichester, U.K.
- Brown, H., and Prescott, R. 2006. *Applied Mixed Models in Medicine*. John Wiley & Sons, Chichester, UK.
- Cipriani, A., Higgins, J. P. T., Geddes, J. R., and Salanti, G. 2013. Conceptual and technical challenges in network meta-analysis. *Ann. Intern. Med.* 159:130-137.
- Denis, J.-B., Piepho, H.-P., and van Eeuwijk, F. A. 1997. Modelling expectation and variance for genotype by environment data. *Heredity* 79:162-171.
- Dias, S., Sutton, A. J., Ades, A. E., and Welton, N. J. 2013a. Evidence synthesis for decision making 2: A generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Med. Decis. Making* 33:607-617.
- Dias, S., Welton, N. J., Sutton, A. J., Caldwell, D. M., Lu, G., and Ades, A. E. 2013b. Evidence synthesis for decision making 4: Inconsistency in networks of evidence based on randomized controlled trials. *Med. Decis. Making* 33:641-656.
- Donegan, S., Williamson, P., D'Allesando, U., and Smith, C. T. 2013. Assessing key assumptions of network meta-analysis: A review of methods. *Res. Synth. Methods* 4:291-323.
- Gauch, H. G., Jr. 1988. Model selection and validation for yield trials with interaction. *Biometrics* 44:705-715.
- Gauch, H. G., Jr., Piepho, H.-P., and Annicchiarico, P. 2008. Statistical analysis of yield trials by AMMI and GGE: Further considerations. *Crop Sci.* 48:866-889.
- Glesser, L. J., and Olkin, I. 2009. Stochastically dependent effect sizes. Pages 357-376 in: *The Handbook of Research Synthesis and Meta-Analysis*. H. Cooper, L. V. Hedges, and J. C. Valentine, eds. Russell Sage Foundation Publishing, New York.
- Greco, T., Edefonti, V., Biondi-Zoccai, G., Decarli, A., Gasparini, M., Zangrillo, A., and Landoni, G. 2015. A multilevel approach to network meta-analysis within a frequentist framework. *Contemp. Clin. Trials* 42:51-59.
- Higgins, J. P. T., Jackson, D., Barrett, J. K., Lu, G., Ades, A. E., and White, I. R. 2012. Consistency and inconsistency in network meta-analysis: Concepts and models for multi-arm studies. *Res. Synth. Methods* 3:98-110.
- Higgins, J. P. T., Thompson, S. G., and Spiegelhalter, D. J. 2009. A re-evaluation of random-effects meta-analysis. *J. R. Stat. Soc. A* 172:137-159.
- Hoaglin, D. C., Hawkins, N., Jansen, J. P., Scott, D. A., Itzler, R., Cappelleri, J. C., Boersma, C., Thompson, D., Larholt, K. M., Diaz, M., and Barrett, A. 2011. Conducting indirect-treatment-comparison and network-meta-analysis studies: Report of the SPOR task force on indirect treatment comparisons good research practices, part 2. *Value Health* 14:429-437.
- Jackson, D., White, I. R., and Riley, R. D. 2013. A matrix-based method of moments for fitting the multivariate random effects model for meta-analysis and meta-regression. *Biom. J.* 55:231-245.
- Jackson, D., and Riley, R. D. 2014. A refined method for multivariate meta-analysis and meta-regression. *Stat. Med.* 33:541-554.
- Jones, B., Roger, J., Lane, P., Lawton, A., Fletcher, C., Cappelleri, J., Tate, H., and Moneuse, P. 2011. Statistical approaches for conducting network meta-analysis in drug development. *Pharm. Stat.* 10:523-531.
- Krahn, U., Binder, H., and König, J. 2013. A graphical tool for locating inconsistency in network meta-analysis. *BMC Med. Res. Methodol.* 13:35.
- Lee, A. W. 2014. Review of mixed treatment comparisons in published systematic reviews show marked increase since 2009. *J. Clin. Epidemiol.* 67:138-143.
- Lipsey, M. W., and Wilson, D. B. 2001. *Practical Meta-Analysis*. Sage Publications Inc., Thousand Oaks, CA.
- Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., and Schabenberger, O. 2006. *SAS for Mixed Models*, 2nd Ed. SAS Institute, Cary, NC.
- Lu, G., and Ades, A. E. 2004. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat. Med.* 23:3105-3124.
- Lu, G., and Ades, A. E. 2006. Assessing evidence consistency in mixed treatment comparisons. *J. Am. Stat. Assoc.* 101:447-459.
- Lu, G., Welton, J. J., Higgins, J. P. T., White, I. R., and Ades, A. E. 2011. Linear inference for mixed treatment comparison meta-analysis: A two-stage approach. *Res. Synth. Methods* 2:43-60.
- Lumley, T. 2002. Network meta-analysis for indirect treatment comparisons. *Stat. Med.* 21:2313-2324.
- Madden, L. V., Bradley, C. A., Dalla Lana da Silva, F., and Paul, P. A. 2014. Meta-analysis of 19 years of fungicide trials for the control of Fusarium head blight of wheat. Pages 17-18 in: *Proceedings of the 2014 National Fusarium Head Blight Forum*. U.S. Wheat & Barley Scab Initiative, East Lansing, MI.
- Madden, L. V., and Paul, P. A. 2011. Meta-analysis for evidence synthesis in plant pathology: An overview. *Phytopathology* 101:16-30.
- Makowski, D., Vincent, A., Pautasso, M., Stancanelli, G., and Rafoss, T. 2014. Comparison of statistical models in a meta-analysis of fungicide treatments for the control of citrus black spot caused by *Phyllosticta citricarpa*. *Eur. J. Plant Pathol.* 139:79-94.
- Malosetti, M., Ribaut, J. M., and van Eeuwijk, F. A. V. 2013. The statistical analysis of multi-environment data: Modeling genotype-by-environment interaction and its genetic basis. *Front. Physiol.* 4:1-17.
- Meyer, K. 2009. Factor-analytic models for genotype x environment type problems and structured covariance matrices. *Genet. Sel. Evol.* 41:21.
- Möhring, J., and Piepho, H.-P. 2009. Comparison of weighting in two-stage analysis of plant breeding trials. *Crop Sci.* 49:1977-1988.
- Nelson, M. E., Gent, D. H., and Grove, G. G. 2015. Meta-analysis reveals a critical period for management of powdery mildew on hop cones. *Plant Dis.* 99:632-640.
- Ng, M. P., and Grunwald, G. K. 1997. Nonlinear regression analysis of the joint-regression model. *Biometrics* 53:1366-1372.
- Ngugi, H. K., Esker, P. D., and Scherm, H. 2011a. Meta-analysis to determine the effects of plant disease management measures: Review and case studies on soybean and apple. *Phytopathology* 101:31-41.
- Ngugi, H. K., Lehman, B. L., and Madden, L. V. 2011b. Multiple treatment meta-analysis of products evaluated for control of fire blight in the Eastern United States. *Phytopathology* 101:512-522.
- Ohlssen, D., Price, K. L., Xia, H. A., Hong, H., Kerman, J., Fu, H., Quartey, G., Heilmann, C. R., Ma, H., and Carlin, B. P. 2014. Guidance on the implementation and reporting of a drug safety Bayesian network meta-analysis. *Pharm. Stat.* 13:55-70.
- Paul, P. A., Hershman, D. E., McMullen, M. P., and Madden, L. V. 2010. Meta-analysis of the effects of triazole-based fungicides on wheat yield and test weight as influenced by Fusarium head blight intensity. *Phytopathology* 100:160-171.
- Paul, P. A., Lipps, P. E., Hershman, D. E., McMullen, M. P., Draper, M. A., and Madden, L. V. 2007. A quantitative review of tebuconazole effect on Fusarium head blight and deoxynivalenol content in wheat. *Phytopathology* 97:211-220.
- Paul, P. A., Lipps, P. E., Hershman, D. E., McMullen, M. P., Draper, M. A., and Madden, L. V. 2008. Efficacy of triazole-based fungicides for Fusarium head blight and deoxynivalenol control in wheat: A multivariate meta-analysis. *Phytopathology* 98:999-1011.
- Paul, P. A., Lipps, P. E., and Madden, L. V. 2006. Meta-analysis of regression coefficients for the relationship between Fusarium head blight and deoxynivalenol content of wheat. *Phytopathology* 96:951-961.

- Paul, P. A., Madden, L. V., Bradley, C. A., Robertson, A. E., Munkvold, G. P., Shaner, G., Wise, K. A., Malvick, D. K., Allen, T., Grybauskas, A., Vincelli, P., and Esker, P. 2011. Meta-analysis of yield response of hybrid field corn to foliar fungicides in the U.S. Corn Belt. *Phytopathology* 101:1122-1132.
- Philibert, A., Loyce, C., and Makowski, D. 2012. Assessment of the quality of the meta-analysis in agronomy. *Agric. Ecosyst. Environ.* 148:72-82.
- Piepho, H.-P. 1997. Analyzing genotype-environment data by mixed models with multiplicative effects. *Biometrics* 53:761-766.
- Piepho, H.-P. 1999. Stability analysis using the SAS system. *Agron. J.* 91: 154-160.
- Piepho, H.-P. 2014. Network meta-analysis made easy: Detection of inconsistency using factorial analysis-of-variance models. *BMC Med. Res. Methodol.* 14:61.
- Piepho, H.-P., Madden, L. V., and Williams, E. R. 2015. Multiplicative interaction in network meta-analysis. *Stat. Med.* 34:582-594.
- Piepho, H.-P., Möhring, J., Schulz-Streeck, T., and Ogutu, J. O. 2012a. A stage-wise approach for the analysis of multi-environment trials. *Biom. J.* 54:844-860.
- Piepho, H.-P., Williams, E. R., and Madden, L. V. 2012b. The use of two-way linear mixed models in multi-treatment meta-analysis. *Biometrics* 68: 1269-1277.
- Pinheiro, J. C., and Bates, D. M. 1995. Approximations to the log-likelihood function in the nonlinear mixed effects model. *J. Comput. Graph. Stat.* 4:12-35.
- Riley, R. D. 2009. Multivariate meta-analysis: The effect of ignoring within-study correlation. *J. R. Stat. Soc. A* 172:789-811.
- Riley, R. D., Abrams, K. R., Sutton, A. J., Lambert, P. C., and Thompson, J. R. 2007. Bivariate random-effects meta-analysis and the estimation of between-study correlation. *BMC Med. Res. Methodol.* 7:3.
- Riley, R. D., Lambert, P. C., and Abo-Zaid, G. 2010. Meta-analysis of individual participant data: rationale, conduct, and reporting. *Br. Med. J. (Clin. Res. Ed.)* 340:c221.
- Riley, R. D., Lambert, P. C., Staessen, J. A., Wang, J., Gueyffier, F., Thijs, L., and Bouillon, F. 2008. Meta-analysis of continuous outcomes combining individual patient data and aggregate data. *Stat. Med.* 27:1870-1893.
- Salanti, G. 2012. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: Many names, many benefits, many concerns for the next generation evidence synthesis tool. *Res. Synth. Methods* 3:80-97.
- Salanti, G., Ades, A. E., Higgins, J. P. T., and Ioannidis, J. P. A. 2008. Evaluation of networks of randomized trials. *Stat. Methods Med. Res.* 17:279-301.
- Salanti, G., and Schmid, C. H. 2012. Research synthesis methods special issue on network meta-analysis: introduction from the editors. *Res. Synth. Methods* 3:69-70.
- Salgado, J. D., Madden, L. V., and Paul, P. A. 2015. Quantifying the effects of Fusarium head blight on grain yield and test weight in soft red winter wheat. *Phytopathology* 105:295-306.
- Schabenberger, O., and Pierce, F. J. 2002. *Contemporary Statistical Models for the Plant and Soil Sciences*. CRC Press, Boca Raton, FL.
- Scherer, H., Thomas, C. S., Garrett, K. A., and Olsen, J. M. 2014. Meta-analysis and other approaches for synthesizing structured and unstructured data in plant pathology. *Annu. Rev. Phytopathol.* 52:453-476.
- Schmid, C. H., Lau, J., Mcintosh, M. W., and Cappelleri, J. C. 1998. An empirical study of the effect of the control rate as a predictor of treatment efficacy in meta-analysis of clinical trials. *Stat. Med.* 17:1923-1942.
- Searle, S. R. 1987. *Linear Models for Unbalanced Data*. John Wiley & Sons, New York.
- Searle, S. R., Casella, G., and McCulloch, C. E. 1992. *Variance Components*. John Wiley & Sons, New York.
- Senn, S. 2010. Hans van Houwelingen and the art of summing up. *Biom. J.* 52: 85-94.
- Senn, S., Gavini, F., Maguez, D., and Scheen, A. 2013. Issues in performing network meta-analysis. *Stat. Methods Med. Res.* 22:169-189.
- Sharp, S. J., and Thompson, S. G. 2000. Analyzing the relationship between treatment effect and underlying risk in meta-analysis: comparison and development approaches. *Stat. Med.* 19:3251-3274.
- Smith, A. B., Cullis, B. R., and Thompson, R. 2005. The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. *J. Agric. Sci.* 143:449-462.
- Stijnen, T., Hamza, T. H., and Özdemir, P. 2010. Random effects meta-analysis of event outcome in the framework of the generalized linear mixed model with applications in sparse data. *Stat. Med.* 29:3046-3067.
- Tu, Y.-K. 2014. Use of generalized linear mixed models for network meta-analysis. *Med. Decis. Making* 34:911-918.
- Tu, Y.-K. 2015. Linear mixed model approach to network meta-analysis for continuous outcomes in periodontal research. *J. Clin. Periodontol.* 42: 204-212.
- van Houwelingen, H. C., Arends, L. R., and Stijnen, T. 2002. Advanced methods in meta-analysis: Multivariate approach and meta-regression. *Stat. Med.* 21:589-624.
- White, I. R., Barrett, J. K., Jackson, D., and Higgins, J. P. T. 2012. Consistency and inconsistency in network meta-analysis: Model estimation using multivariate meta-regression. *Res. Synth. Methods* 3:111-125.
- Whitehead, A. 2002. *Meta-Analysis of Controlled Clinical Trials*. John Wiley & Sons, West Sussex, England.
- Willyerd, K. T., Li, C., Madden, L. V., Bradley, C. A., Bergstrom, G. C., Sweets, L. E., McMullen, M., Ransom, J. K., Grybauskas, A., Osborne, L., Wegulo, S. N., Hershman, D. E., Wise, K., Bockus, W. W., Groth, D., Dill-Macky, R., Milus, E., Esker, P. D., Waxman, K. D., Adey, E. A., Ebelhar, S. E., Young, B. G., and Paul, P. A. 2012. Efficacy and stability of integrating fungicide and cultivar resistance to manage Fusarium head blight and deoxynivalenol in wheat. *Plant Dis.* 96:957-967.
- Yan, W., Hunt, L. A., Sheng, Q., and Szlavnick, Z. 2000. Cultivar evaluation and mega-environment investigation based on the GGE biplot. *Crop Sci.* 40: 597-605.