

New Applications of Statistical Tools in Plant Pathology

Meta-Analysis in Plant Pathology: Synthesizing Research Results

M. S. Rosenberg, K. A. Garrett, Z. Su, and R. L. Bowden

First author: School of Life Sciences, Arizona State University, Tempe 85287; second and third authors: Department of Plant Pathology, Kansas State University, Manhattan 66506; and fourth author: U.S. Department of Agriculture-Agricultural Research Service, Plant Science and Entomology Research Unit, Manhattan, KS 66506.

Current address of Z. Su: Statistical and Data Analysis Center, Harvard School of Public Health, Boston, MA 02115.

Accepted for publication 16 May 2004.

ABSTRACT

Rosenberg, M. S., Garrett, K. A., Su, Z., and Bowden, R. L. 2004. Meta-analysis in plant pathology: Synthesizing research results. *Phytopathology* 94:1013-1017.

Meta-analysis is a set of statistical procedures for synthesizing research results from a number of different studies. An estimate of a statistical effect, such as the difference in disease severity for plants with or without a management treatment, is collected from each study along with a measure of the variance of the estimate of the effect. Combining results from different studies will generally result in increased statistical power so that

it is easier to detect small effects. Combining results from different studies may also make it possible to compare the size of the effect as a function of other predictor variables such as geographic region or pathogen species. We present a review of the basic methodology for meta-analysis. We also present an example of meta-analysis of the relationship between disease severity and yield loss for foliar wheat diseases, based on data collected from a decade of fungicide and nematicide test results.

Additional keyword: plant productivity.

Research falls into two broad categories: primary research, in which an investigator examines a particular phenomenon, and research synthesis, in which an investigator reviews and summarizes primary research. In its broadest sense, meta-analysis refers to any quantitative review of the primary literature (13). For the most part, however, meta-analysis refers to a specific set of statistical procedures that are used to summarize primary studies by estimating measures of overall effect, consistency and homogeneity among the studies, and whether the effect differs significantly from a null expectation. As with any research, meta-analysis contains a wide variety of steps from planning to data collection to analysis to interpretation. We will give a brief overview and example; for more information on all aspects of meta-analysis, we recommend Hedges and Olkin (17), Cooper and Hedges (11), Normand (23), Rosenberg et al. (28), and Arthur et al. (2), as well as Hunt (19) for a historical perspective and overview.

The analytical part of a meta-analysis consists of two main parts: estimation of effect size from individual studies and estimation of summary statistics. An effect is a statistical estimate of the degree to which a phenomenon is present in a study (10). For a specific phenomenon there may be many ways of measuring effect. For example, the effect of a biological control or fungicide compared with no treatment on the amount of disease could be measured as the difference in disease severity between plants treated and untreated, or the difference in disease incidence, or the ratio of disease severity, or the ratio of disease incidence. Each of these is a valid way of measuring effect and their magnitude is known as the effect size. In a meta-analysis, one chooses

a single effect measure and estimates it for all studies; this allows us to compare and combine the results of these studies directly.

Effect size can be estimated in any number of ways; specific choice of a measure often depends on both the specific questions being asked and the nature of the data. Most meta-analyses use one of a number of common measures of effect size, but it is possible, and sometimes preferable, to use an *ad-hoc* measure (25,26). We will describe briefly the most common effect measures below; for a broader overview see Rosenberg et al. (28).

In plant pathology, there are a number of questions that can be addressed using meta-analysis. Likely goals for meta-analysis include estimates of the effects of management techniques on disease abundance, the effects of disease abundance on plant characteristics such as yield, and the effects of type of resistance and form of resistance deployment on durability of resistance. Summaries of the reported host range of pathogens (available online from the USDA-ARS Systematic Botany and Mycology Laboratory website and the Virus Identification Data Exchange Project) also offer an opportunity for analyzing how host or pathogen characteristics affect host range and types; for example, Mitchell and Power (20) found that the number of reported pathogen species was lower for introduced compared to native host plant species, consistent with the hypothesis that invasive plant species experience some release from pathogen pressure. In addition to the potential for drawing data from journals such as *Phytopathology* and *Plant Disease*, the annual series *Fungicide and Nematicide Tests* (FNT) and *Biological and Cultural Tests* (BCT) offer an abundance of data. These publications include data relevant to questions about the efficacy of pesticides in terms of different timings, rates, and adjuvants, about biological and cultural controls, and about the relationship between disease severity and yield. FNT and BCT also offer the advantage that results that are not statistically significant are more likely not to be excluded from publication. Bias in the acceptance of studies reporting different levels of statistical significance is a concern; for example, in a

Corresponding author: K. A. Garrett; E-mail address: kgarrett@ksu.edu

Publication no. P-2004-0719-04O

© 2004 The American Phytopathological Society

study of ecological meta-analyses, Murtaugh (21) found a relationship between effect size and the scientific citation rating of the journal in which the result was published. Other challenges for meta-analyses of data in plant pathology include the fact that multiple diseases are often present on the same experimental plants, inconsistent reporting of the sampling variation, and the inconsistent use of standard cultivars that would allow for easier comparisons across years and regions. But meta-analysis may be an important tool for more formally synthesizing data from FNT experiments for evaluation of fungicide or nematicide registration.

Another potential application of meta-analysis in plant pathology is in the analysis of accumulated data sets within one or a few research programs. Smiley and Patterson (31) and Olkin and Shaw (24) give examples of studies in which a small number of experiments were not sufficient to provide statistical support for an observed trend. Smiley and Patterson (31) found that combining several estimates of the effect of wheat seed treatments provided adequate statistical power to demonstrate an effect of the treatments on yield. Olkin and Shaw (24) considered biological and chemical control of spider mites in strawberries. All 10 biocontrol studies that they included in their meta-analysis reported beneficial effects that were not, however, statistically significant. The meta-analysis also showed an effect of biological control and the increased statistical power gave a statistically significant result. Another agricultural meta-analysis considered gains in agricultural productivity; Grandillo et al. (14) examined the genetic improvement of processing tomatoes over 20 years in terms of yield and total soluble solids content. Inclusion of common check varieties in the experiments allowed them to account for environmental factors. Garrett et al. (K. A. Garrett, L. N. Zúñiga, E. Roncal, G. A. Forbes, C. C. Mundt, Z. Su, and R. J. Nelson, *unpublished data*) analyzed results from several studies of the effects of cultivar mixtures on potato late blight. Combining the results allowed consideration of how the effect varied as a function of the degree of seasonality of the different locations studied.

Additionally, there have been a number of relevant ecological meta-analyses in recent years. Borowicz (6) found that inoculation with arbuscular mycorrhizal fungi generally had a large negative effect on growth of pathogens. Hawkes and Sullivan (15) concluded that overcompensation in plants in response to herbivory was more likely in high resource environments for monocots and in low resource environments for dicots. Searles et al. (30) found that plant morphological characteristics and photosynthetic processes showed little response in studies of simulated stratospheric ozone depletion.

Meta-analysis methods. A first step in meta-analysis is determination of the effect to be analyzed over studies. The simplest measure of effect is the correlation coefficient, r . It is perhaps the measure used most widely for effect size, particularly in the social sciences. One advantage of r is that it has fewer data requirements and is easier to calculate than many alternate measures. Also, it is possible to transform a variety of statistics (including normal variates, t scores, and χ^2) into correlation coefficients (28,29). Because r does not have desirable statistical properties, it is first transformed using Fisher's z transformation (29):

$$z = \frac{1}{2} \ln \left(\frac{1+r}{1-r} \right) \quad (1)$$

The z transform will result in a value ranging from positive to negative infinity, with a value of zero indicating no effect; furthermore, the variance for z is estimated simply as

$$v_z = \frac{1}{n-3} \quad (2)$$

where n is the number of studies.

Another common measure is the standardized mean difference. There are many variations on this effect measure (28), but the one

most often used with the best statistical properties is known as Hedges' d . Given means (\bar{X}^E, \bar{X}^C), standard deviations (s^E, s^C), and sample sizes (N^E, N^C) for an experiment and control group, respectively, the standardized mean difference can be calculated as

$$d = \frac{(\bar{X}^E - \bar{X}^C)}{S} J \quad (3)$$

where S is a measure of the pooled variance,

$$S = \sqrt{\frac{(N^C - 1)(s^C)^2 + (N^E - 1)(s^E)^2}{N^C + N^E - 2}} \quad (4)$$

and J is a correction for small sample sizes,

$$J = 1 - \frac{3}{4(N^C + N^E - 2)} \quad (5)$$

As sample sizes increase, J approaches 1 asymptotically. The variance for d is estimated by

$$v_d = \frac{N^C + N^E}{N^C N^E} + \frac{d^2}{2(N^C + N^E)} \quad (6)$$

There are many other potential measures of effect size. A common form of data, particularly in the medical literature, is the two by two contingency table. Data of this sort lends itself to measures of effect size such as odds ratios and relative rates (23,28). *Ad-hoc* measures can be advantageous, because they may fit the specific question being addressed and allow one to model a process (25,26). However, they can be much more difficult to deal with, particularly since it is necessary to estimate the variance of the metric.

When performing a meta-analysis, it may be difficult to determine the appropriate measure of variance to include for some studies. Ideally the original data might be available for analysis, but this will not generally be the case. In publications like those in FNT, the mean squared error (MSE) or least significant difference (LSD) is sometimes included and these can be used to calculate estimates of variability. If the MSE is reported, then the reciprocal of the square root of the MSE would be used as the weight in the meta-analysis; if the LSD is reported, then the MSE could be computed from the LSD. For example, in a randomized complete block design, $LSD = t_{\alpha/2} \sqrt{2MSE/b}$, where b is the number of blocks and $t_{\alpha/2}$ follows the t distribution with significance level α and degrees of freedom equal to $(b-1)(k-1)$. Here, k is the number of treatments in the study.

Since the LSD is calculated over all treatments, it may not be an adequate measure of variation for all meta-analysis purposes. When the responses to some treatments are much greater than others, the variance will often also be greater for these treatments compared with that of others. In such a case, it would be most useful if the standard error for each treatment mean were reported.

Once an effect size is estimated for each study, the next step is to summarize these results from the individual studies to gain an understanding of the overall effect. In the following, E_i indicates the effect size for the i th study, regardless of which effect measure is used, and v_i indicates its variance. In meta-analysis, studies are weighted differentially depending on their accuracy. The simplest and least controversial method of estimating their accuracy is variance; studies with low variance are given high weight and those with high variance are given low weight. Specifically, the weight for the i th study is the inverse of its variance,

$$w_i = \frac{1}{v_i} \quad (7)$$

The grand measure of effect is simply a weighted average of the individual effect sizes,

$$\bar{E} = \frac{\sum w_i E_i}{\sum w_i} \quad (8)$$

with variance

$$s_E^2 = \frac{1}{\sum w_i} \quad (9)$$

This is the average effect for all studies. Confidence intervals can be created to test whether the mean differs from the hypothesis of no effect using standard distributions or bootstrapping (1). To test whether the sample effect sizes are themselves homogeneous (from a single population), one uses Q statistics, a form of weighted sums-of-squares. The total heterogeneity in a sample (17) is

$$Q_T = \sum w_i (E_i - \bar{E})^2 \quad (10)$$

Q_T can be tested against a χ^2 distribution with $n - 1$ degrees of freedom. This tests whether the effect sizes for all of the studies are equal; a significant result indicates that the variance is greater than expected due to sampling error (12). Simple calculation of \bar{E} assumes that all of the studies come from a single homogeneous sample. Often variation in the effect size may be explainable by additional variables. The simplest ways to evaluate this follow the framework of analysis of variance (ANOVA) and linear regression.

Under an ANOVA-like structure, one subdivides the studies into a number of categories and wishes to know whether these groups (i) differ in their respective mean effect size, and (ii) show significant heterogeneity of effect among groups relative to within groups. Mean effect sizes (and their variances) for each group (\bar{E}_j) and the grand mean for all groups (\bar{E}) are determined simply as weighted averages as above. Heterogeneity within an individual group j is calculated as above,

$$Q_{w_j} = \sum w_{ij} (E_{ij} - \bar{E}_j)^2 \quad (11)$$

where E_{ij} is the effect size of the i th study in the j th group. This is tested against a χ^2 with $n_j - 1$ degrees of freedom, where n_j is the number of studies in the j th group. The total heterogeneity statistic (Q_T) can be divided into heterogeneity within (Q_W) and among groups (Q_B),

$$Q_T = Q_B + Q_W \quad (12)$$

where

$$Q_W = \sum \sum w_{ij} (E_{ij} - \bar{E}_j)^2 = \sum Q_{w_j} \quad (13)$$

and

$$Q_B = \sum \sum w_{ij} (\bar{E}_j - \bar{E})^2 \quad (14)$$

These have $n - m$ and $m - 1$ degrees of freedom, respectively, where m is the number of groups. A significant Q_B indicates that a significant portion of the observed heterogeneity can be explained by subdividing the studies into the categories.

Instead of dividing the studies into set groups, one may wish to test whether the effect sizes of individual studies are dependent on another (independent) variable, i.e., linear regression. In meta-analysis, linear regression uses the standard model to explain the dependent variable,

$$E_i = b_0 + b_1 X_i + \varepsilon \quad (15)$$

where b_0 and b_1 are the intercept and slope and X is the independent variable. The slope and intercept are determined using weighted linear regression:

$$b_1 = \frac{\sum w_i X_i E_i - \frac{\sum w_i X_i \sum w_i E_i}{\sum w_i}}{\sum w_i X_i^2 - \frac{(\sum w_i X_i)^2}{\sum w_i}} \quad (16)$$

and

$$b_0 = \frac{\sum w_i E_i - b_1 \sum w_i X_i}{\sum w_i} \quad (17)$$

Standard errors for b_0 and b_1 can be found in Hedges and Olkin (17) and Rosenberg et al. (28) and can be used to test whether they differ from zero. As above, Q_T can be divided into heterogeneity due to the model, Q_M , and error heterogeneity, Q_E (equivalent to Q_B and Q_W , respectively). In this case,

$$Q_M = \frac{b_1^2}{s_b^2} \quad (18)$$

and Q_E is determined most easily as $Q_E = Q_T - Q_M$. Q_M and Q_E are also tested against a χ^2 with 1 and $n - 2$ degrees of freedom, respectively. A significant Q_M indicates that a substantial amount of the observed heterogeneity is explained by the regression model. A significant Q_E indicates that there is substantial unexplained variation.

More complicated statistical models (e.g., ANCOVA or two-way ANOVA) are possible in meta-analysis; see Rosenberg et al. (28) for details.

The above calculations are all based on a fixed-effects meta-analysis model. A fixed-effects model assumes that there is a single true effect in all studies (or for each group of studies) and that any observed variation is due to sampling error (16). An alternate model, the random-effects model (27), assumes that there is an average effect with a certain degree of variation around this mean; observed heterogeneity is due to this variation and is not simply sampling error. To incorporate a random-effects model in a meta-analysis, one must simply recalculate the individual study weights. Recall that in a fixed-effects model meta-analysis, the weight of each study is simply the inverse of its variance. For a random-effects model,

$$w_i = \frac{1}{v_i + \sigma_{pooled}^2} \quad (19)$$

The weight of each study is equal to the inverse of the study's variance plus an estimate of the pooled variance. The formula for pooled variance varies depending on the type of statistical model being tested. When calculating a simple grand mean, the pooled variance is

$$\sigma_{pooled}^2 = \frac{Q_T - (n-1)}{\sum w_i - \frac{\sum w_i^2}{\sum w_i}} \quad (20)$$

where Q_T is the total heterogeneity determined from a fixed-effects model meta-analysis. Similarly, when dividing the studies into fixed categories,

$$\sigma_{pooled}^2 = \frac{Q_W - (n-m)}{\left(\sum w_{ij} - \frac{\sum w_{ij}^2}{\sum w_{ij}} \right)} \quad (21)$$

Formulas for the pooled variance in regression models and other complicated statistical models can be found in Rosenberg et al. (28). Once the new weights are determined for a random-effects model, all calculations proceed exactly as for the fixed-effects model. It should be noted that the error heterogeneity (Q_E) will never be significant in a random-effects model since the purpose of the model is to explain this error as due to the nature of the studies and not to sampling.

There have been a number of software packages produced for meta-analysis, ranging from commercial software to freeware, including one by an author of this paper, MetaWin (28). Analyses can also be performed within statistical packages such as SAS (SAS Institute, Cary, NC) (2). Unfortunately, we know of no

recent comparative reviews of the available software and older published reviews too often refer to out-dated software no longer available. To those interested, we would suggest performing a thorough web search to find information on the current availability of meta-analytic software.

Meta-analysis of the relationship between foliar disease severity and yield loss. We performed a meta-analysis of the relationship between foliar disease severity and yield loss in wheat. The goal of the analysis was to estimate this relationship for a number of wheat diseases and compare the estimates between diseases and, as the data set allowed, between geographic regions. For this analysis, we used volumes 45 to 54 (1990 to 1999) of FNT. There were a total of 175 reports on wheat yield loss in this 10-year period, 25 of which dealt with a single foliar disease. To make results more directly comparable from one study to another, we used the following criteria for inclusion in our meta-analysis. (i) At least one treatment had disease severity near zero for use in estimation of the relative yield loss. (ii) Disease severity was measured as the percentage of diseased area of the flag leaf. (iii) Yield was reported as volume per cropping area or sufficient information was given to calculate volume per area. (iv) Disease severity was estimated at the soft dough stage. Only five studies in the 10-year period met all four of these criteria: two studies of leaf rust (7,8) and three studies of tan spot (3–5).

Typical studies in FNT report yield and a measure of disease severity or pathogen abundance for each of several fungicide or nematicide treatments. Each treatment is generally replicated, but only the means are reported. Measures of variation associated with the results for each treatment are not usually reported, although an overall LSD and/or standard deviation may be reported. Because the studies in FNT do not generally analyze the relationship between disease severity and yield directly, we estimated this relationship for each of the studies using the reported data. Within each of the four studies included in our meta-analysis, the GLM (general linear models) procedure in SAS (SAS Institute) was used to fit the simple linear regression model $Y = \beta_0 + \beta_1 X + \varepsilon$ to the mean values reported, where X is the percent disease severity and Y is the yield. The response of yield to disease severity has been modeled to take into account the asymptotic nature of the relationship (9), but over the range of disease

severity in our study, a linear approximation appeared reasonable. (Hughes [18] has described the potential for bias when calculating yield responses based on mean responses. If the response of yield to disease severity is not a straight line, using the mean severity and mean loss to estimate the relationship between yield and severity will produce a bias. This might be taken into account in a more elaborate analysis.)

The relative yield loss percentage of Z was obtained for each study as follows. First, the intercept of the regression line for each study was determined using the above regression model. The intercept β_0 was taken as the estimate of the yield potential in the absence of disease. Then Z was set equal to the yield potential β_0 minus the yield for a given disease severity and then divided by the yield potential, $Z = (\beta_0 - Y)/\beta_0$. The yield loss values (Z) were fit in a second linear regression to estimate the relationship between relative yield loss and percentage of disease severity with the intercept set at 0. The slope from this second regression analysis and the associated standard error were used in the meta-analysis.

For the meta-analysis, the slope estimates from each of the four studies were included in a weighted least squares regression analysis where the weights were the reciprocals of the standard error of the slope estimates. This analysis provides an overall estimate of yield loss for foliar wheat pathogens. We also compared the slope for the two diseases in our data set, leaf rust and tan spot, though the level of statistical power for such a comparison was low since there were such a small number of replicates for each disease.

The weighted overall slope estimate based on the five studies was 0.309 (Fig. 1), with standard error 0.055 and 95% confidence interval (0.20, 0.36). For each percent increase in disease severity on the flag leaves of wheat at the soft dough stage, the estimated relative yield loss increased by 0.309%. Q statistics were used to test whether the slopes for all the studies were equal. In this low-power study, the null hypothesis of equality was not rejected: $Q_T = 1.92$ ($P = 0.59$). Tan spot and leaf rust were also compared using Q statistics and the difference between the diseases was not statistically significant ($Q_B = 0.15$, $Q_W = 1.76$, $3Q_B/Q_W = 0.26$, $P = 0.65$).

This meta-analysis is limited by the small number of data sets that were finally included in the analysis, but it serves an illustrative purpose. Our criteria could have been relaxed to allow studies that included another disease, but only at very low levels. We also could have considered allowing evaluations of disease severity over more than just the flag leaf or at different growth stages, or even using different systems such as the stepwise scale often used for powdery mildew, although this would have made the comparisons less direct. Of course, our collection of data sets might have been expanded if we had asked a question more directly related to the goals of FNT; for example, we might have compared the effect of different fungicide formulations or timings on disease severity and yield.

Our final data collection could also have been expanded if more information were typically included in FNT studies. It may be helpful for researchers who publish in FNT, or any forum, to consider how easily their studies could be incorporated into larger future studies. Standardized rating systems make synthesis more straightforward, even if only within a particular host–pathogen system. Clearly, reporting the variance associated with estimates and sample sizes will allow results to be included in weighted analyses. For online publications, it would be desirable to make the complete data set available, perhaps after an appropriate embargo period so that the authors can perform any other desired analyses. The accumulated data sets would be a valuable resource, as illustrated by the Global Population Dynamics Database maintained by the U.K. Centre for Population Biology and the U.S. National Center for Ecological Analysis and Synthesis (22).

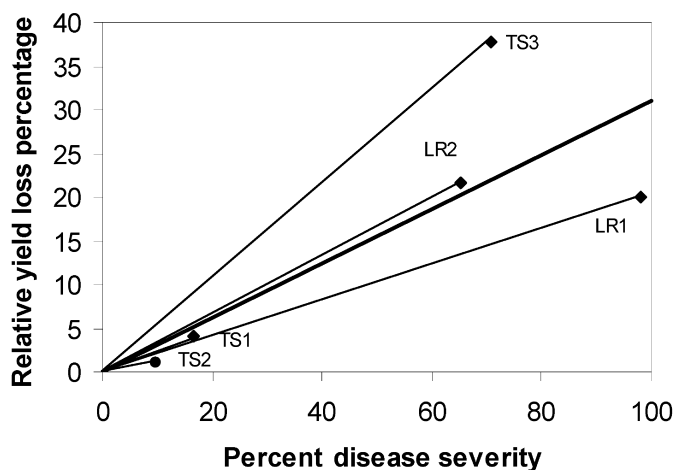


Fig. 1. A meta-analysis of the relative yield loss percentage as a function of the percent wheat flag leaf infected at the soft dough stage. The relative yield loss percentage was the percent lost compared with the yield for nearly disease-free controls in each study. Each point represents the average response for a particular fungicide treatment as reported in a publication, and the horizontal value is the maximum percent disease severity observed in that study. Estimates for leaf rust were based on two publications: LR1 (7) and LR2 (8). Estimates for tan spot were based on three publications: TS1 (3), TS2 (4), and TS3 (5).

ACKNOWLEDGMENTS

We thank W. W. Bockus, T. C. Todd, and *Phytopathology* reviewers for comments that improved this manuscript and S. P. Dendy and A. Berry for help assembling the meta-analysis. This work was supported in part by NSF grants DEB-0130692, EPS-0236913 with matching funds from the Kansas Technology Enterprise Corporation, and EPS-9874732 with matching support from the State of Kansas, and by USDA Grant 2002-34103-11746. This is Kansas State Experiment Station Contribution 04-259-J.

LITERATURE CITED

1. Adams, D. C., Gurevitch, J., and Rosenberg, M. S. 1997. Resampling tests for meta-analysis of ecological data. *Ecology* 78:1277-1283.
2. Arthur, W., Jr., Bennett, W., Jr., and Huffcutt, A. I. 2001. *Conducting Meta-Analysis Using SAS*. Lawrence Erlbaum Associates, Mahwah, NJ.
3. Bockus, W. W. 1995. Effect of timing of foliar fungicides on production of large seed of winter wheat. *Fung. Nemat. Tests* 50:221.
4. Bockus, W. W. 1995. Evaluation of foliar fungicides on winter wheat for control of tan spot. *Fung. Nemat. Tests* 50:222.
5. Bockus, W. W. 1996. Effect of foliar fungicides on production of grain and large seed of winter wheat. *Fung. Nemat. Tests* 51:200.
6. Borowicz, V. A. 2001. Do arbuscular mycorrhizal fungi alter plant-pathogen relations? *Ecology* 82:3057-3068.
7. Bowden, R. L. 1992. Evaluation of fungicides for control of wheat leaf rust in western Kansas. *Fung. Nemat. Tests* 47:182.
8. Bowden, R. L. 1993. Evaluation of fungicides for control of wheat leaf rust in western Kansas. *Fung. Nemat. Tests* 48:226.
9. Campbell, C. L., and Madden, L. V. 1990. *Introduction to Plant Disease Epidemiology*. Wiley, New York.
10. Cohen, J. 1969. *Statistical Power Analysis for the Behavioral Sciences*. Academic Press, New York.
11. Cooper, H., and Hedges, L. V. 1994. *The Handbook of Research Synthesis*. Russell Sage Foundation, New York.
12. Cooper, H. M. 1998. *Synthesizing Research: A Guide for Literature Reviews*. 3rd ed. Vol. 2. Sage, Beverly Hills, CA.
13. Glass, G. V. 1976. Primary, secondary, and meta-analysis of research. *Educ. Res.* 5:3-8.
14. Grandillo, S., Zamir, D., and Tanksley, S. D. 1999. Genetic improvement of processing tomatoes: A 20 year perspective. *Euphytica* 110:85-97.
15. Hawkes, C. V., and Sullivan, J. J. 2001. The impact of herbivory on plants in different resource conditions: A meta-analysis. *Ecology* 82:2045-2058.
16. Hedges, L. V. 1994. Statistical considerations. Pages 29-38 in: *The Handbook of Research Synthesis*. H. Cooper and L. V. Hedges, eds. Russell Sage Foundation, New York.
17. Hedges, L. V., and Olkin, I. 1985. *Statistical Methods for Meta-Analysis*. Academic Press, San Diego.
18. Hughes, G. 1988. Spatial heterogeneity in crop loss assessment models. *Phytopathology* 78:883-884.
19. Hunt, M. 1997. *How Science Takes Stock: The Story of Meta-Analysis*. Russell Sage Foundation, New York.
20. Mitchell, C. E., and Power, A. G. 2003. Release of invasive plants from fungal and viral pathogens. *Nature* 421:625-627.
21. Murtaugh, P. A. 2002. Journal quality, effect size, and publication bias in meta-analysis. *Ecology* 83:1162-1166.
22. NERC Centre for Population Biology, Imperial College. 1999. *The Global Population Dynamics Database*. <http://www.sw.ic.ac.uk/cpb/cpb/gpdd.html>.
23. Normand, S.-L. 1999. *Tutorial in Biostatistics Meta-Analysis*. Russell Sage Foundation, New York.
24. Olkin, I., and Shaw, D. V. 1995. Meta-analysis and its applications in horticultural science. *HortScience* 30:1343-1348.
25. Osenberg, C. W., Sarnelle, O., and Cooper, S. D. 1997. Effect size in ecological experiments: The application of biological models in meta-analysis. *Am. Naturalist* 150:798-812.
26. Osenberg, C. W., and St. Mary, C. M. 1997. Meta-analysis: Synthesis or statistical subjugation? *Integrative Biol.* 1:37-41.
27. Raudenbush, S. W. 1994. Random effects models. Pages 301-321 in: *The Handbook of Research Synthesis*. L. V. Hedges, ed. Russell Sage Foundation, New York.
28. Rosenberg, M. S., Adams, D. C., and Gurevitch, J. 2000. *MetaWin. Statistical Software for Meta-Analysis*. Sinauer Associates, Sunderland, MA.
29. Rosenthal, R. 1991. *Meta-Analytic Procedures for Social Research*. Rev. Ed. Sage, Newbury Park.
30. Searles, P. S., Flint, S. D., and Caldwell, M. M. 2001. A meta-analysis of plant field studies simulating stratospheric ozone depletion. *Oecologia* 127:1-10.
31. Smiley, R. W., and Patterson, L. M. 1995. Winter wheat yield and profitability from Dividend and Vitavax seed treatments. *J. Prod. Agric.* 8:350-354.