HAL
open science

Creating a suite of macros for meta-analysis in SAS: A case study in collaboration

Conny Berlin, Michael Branson, Ekkehard Glimm, Tsushung A. Hua, Stephen Senn, James Weir

To cite this version:

Conny Berlin, Michael Branson, Ekkehard Glimm, Tsushung A. Hua, Stephen Senn, et al.. Creating a suite of macros for meta-analysis in SAS: A case study in collaboration. Statistics and Probability Letters, 2011, 81 (7), pp.842. 10.1016/j.spl.2011.02.010 . hal-00746101

HAL Id: hal-00746101

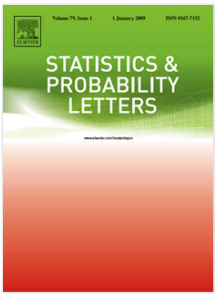
https://hal.science/hal-00746101

Submitted on 27 Oct 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Accepted Manuscript



Creating a suite of macros for meta-analysis in SAS®: A case study in collaboration

Stephen Senn, James Weir, Tsushung A. Hua, Conny Berlin, Michael Branson, Ekkehard Glimm

PII: S0167-7152(11)00048-4
DOI: 10.1016/j.spl.2011.02.010
Reference: STAPRO 5908

To appear in: *Statistics and Probability Letters*

Please cite this article as: Senn, S., Weir, J., Hua, T.A., Berlin, C., Branson, M., Glimm, E., Creating a suite of macros for meta-analysis in SAS®: A case study in collaboration. *Statistics and Probability Letters* (2011), doi:10.1016/j.spl.2011.02.010

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

F
Conny Berlin
Novartis Pharma AG
Drug Safety & Epidemiology
CHBS, WSJ-027.2.097
Novartis Pharma AG
Forum 1
Novartis Campus
CH-4056 Basel
Switzerland
Phone: +41 61 3248736
Fax: +41 61 3248922
conny.berlin@novartis.com

Michael Branson
Translational Sciences
Novartis Pharma AG,
WSJ-210.4.32
CH 4056 Basel
Switzerland
Tel +41 61 324 3407
michael.branson@novartis.com

Ekkehard Glimm
Novartis Pharma AG
Statistical methodology, IIS
CHBS, WSJ-027.1.015
Novartis Pharma AG
Forum 1
Novartis Campus
CH-4056 Basel
Switzerland
Phone: +41 61 3240173
Fax: +41 61 3243039
ekkehard.glimm@novartis.com

Tsushung A Hua
PH, DEV-IIS-CV Biostatistics
USEH, 419-2154
Novartis Pharmaceuticals Corporation
One Health Plaza
East Hanover, NJ 07936-1080
USA
Tel: +1 862 7782629
Fax: +1 973 7816498
tsushung.hua@novartis.com

Stephen Senn
School of Mathematics and Statistics
University of Glasgow
15 University Gardens
Glasgow G12 8QW
UK

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Tel: +44 (0)141 330 5141
Fax: +44 (0)141 330 4814
stephen.senn@glasgow.ac.uk

Jim Weir
Centre for Population Health Sciences,
University of Edinburgh,
Edinburgh,
EH8 9AG
Tel: +44 (0) 131 650 3035
Fax: +44 (0) 131 650 3224
Jim.weir@ed.ac.uk

ACCEPTED MANUSCRIPT

*Manuscript
Click here to view linked References
ACCEPTED MANUSCRIPT

Creating a Suite of Macros for Meta-analysis in SAS®: A Case Study in Collaboration

Stephen Senn, James Weir, Tsushung A Hua, Conny Berlin, Michael Branson and Ekkehard Glimm

Abstract

A series of macros that have been created to perform fixed and random effects meta-analysis in SAS® are described as is the motivation for their creation. These macros are being made freely available on the internet for others to use. The application of the macros is illustrated using an example of trials in pre-eclampsia.

Keywords: Fixed effects; random effects; binary; Galbraith plot; forest plot; funnel plot

1. Introduction

There are various reasons why any pharmaceutical sponsor ought to be interested in meta-analysis. First, it has become increasingly clear that to understand fully what the properties of any pharmaceutical are, whether in terms of efficacy or safety, requires careful formal summaries of the available clinical evidence. Since sponsors wish to understand the properties of their products as fully as possible, use of meta-analysis is appropriate. Second, many regulatory authorities now expect to see such summaries, not as a substitute for successfully conducted individual trials, but as a summary of the evidence from such trials. Third, a company may be interested in using such summaries to assist in planning further studies with different molecules in the same indication. Fourth, it may be desired to make inferences about subgroups of patients for which purpose individual trials may be insufficiently precise. Lastly, many external parties are carrying out such analyses and the need to react to information about its products that others have obtained can be significantly reduced.
Although there are many specialised meta-analysis packages available to do this, they do not necessarily have all the features that were considered desirable. As the main statistical analysis package used for statistical assessment of clinical data in the pharmaceutical industry is SAS®, it was felt that developing a suite of SAS® macros would be logical on grounds of convenience, compatability and consistency. SAS® also has many powerful features that can be exploited. For instance, proc nlmixed® makes the analysis of the sort of non-linear mixed effects models required for random effects for binary data extremely easy. It must also be conceded that SAS® has some disadvantages, including that it requires more work to transfer information between one procedure and another than is the case with object-oriented languages. Nevertheless, on balance, it was considered desirable to develop the set of macros in SAS®.

ACCEPTED MANUSCRIPT

Of course, we are not the first to have produced SAS® code for performing meta-analysis. Such code is, for example, included in the fine tutorial paper on meta-analysis by Normand (Normand, 1999) in *Statistics in Medicine*. In the same journal one can also find a useful tutorial paper by (van Houwelingen et al., 2002). However, both of these, interesting and useful as they are, are unlikely to fit the bill entirely for anyone wishing to do regular meta-analyses. For example, the latter offers a non-standard approach to modelling the main effect of trials, which it treats as random(Senn, 2010). This is not the standard approach used even in a so-called random effects meta-analysis. Also, it has to be conceded that many packages include routines for meta-analysis. For example, Stata® is particularly strong in this respect and one of us, SS, has collaborated to provide a procedure to perform meta-analyses with GenStat®. Nevertheless, there are advantages in having dedicated macros to do exactly the range of analyses one needs in the package one standardly uses to manage data and produce analyses and it is the creation of these macros and their functionality that we describe here.

One of us (SS) had already created a program in Mathcad® to carry out fixed and random effects meta-analysis. However, since Mathcad® has more limited facilities for conventional statistical analysis and is not widely used in the pharmaceutical industry, it was decided to use the Mathcad® program as a basis but create a SAS® equivalent. In addition it was decided to add various other facilities. For example, a paper by Lee(Lee, 1999) provides various simple checks that can be made when meta-analyses of reported binary data are performed. Although in most cases it was anticipated that the meta-analyses would be performed on data generated 'in-house' so that Lee's checks would be in principle at least redundant, it was agreed that it could be important on occasion to analyse externally reported data. Hence a module to perform Lee's checks was to be included.

It was also considered absolutely vital to validate the modules, because health authorities generally require this for any code that is used routinely for submission-relevant analyses.

This article describes the suite of meta-analysis macros that was created and the steps involved in their creation within this specific academic-industry collaboration. It has two purposes. The first is to introduce the macros to others. They are being made freely available on the web at http://www.senns.demon.co.uk/SAS%20Macros/SASMacros.html . The second is that it is felt production of the macros raises some issue of analysis that may be of interest to others.

We have made the macros freely available in 'pre-production form' (that is to say in a form prior to their final versions adopted by the sponsor, which are imbedded into the information technology infrastructure, including links to clinical projects, change tracking etc.). These macros may be used freely, however, all responsibilities regarding validation / code integrity resides with the users – no responsibility is or will be borne by the collaborators or their respective academic/commercial companies. The macros are thus provided as is without any guarantees or warranty.

ACCEPTED MANUSCRIPT

2. Basic philosophy of the software development

For reasons already explained it was decided to develop the software as a series of SAS® macros. The following are a number of choices that could face any person wishing to carry out a meta-analysis of data.

First, it could be decided to carry out an analysis of original data. As an example, consider the analysis of recombinant human erythropoeisi-stimulating agents carried out by Bohlius et al. (Bohlius et al., 2009). Here, enormous efforts were made to obtain original data from various sources. However, when original data are available the analysis will often not require the use of specialist meta-analysis software. For instance, to analyse survival data from a number of trials with a common protocol, an obvious approach is to fit a proportional hazards model with trial as a factor. Indeed that was one of the models considered by Bohlius et al. [2]. In fact, the need for specialist software is likely to be greater where only summary data are available. Accordingly, it is for this more 'classical' situation that the programs were developed. Of course, for certain types of data, it makes no difference whether summaries or original values are used. This is recognised in standard statistical approaches to logistic regression where either the data can be entered as a series of Bernoulli trials, one line for each patient or (if covariates are not being used) as binomial totals for successes, and sample sizes grouped by trial and treatment.

In fact it was decided at the beginning to create a separate module for binary data for which a wide variety of choices for analysis are available.

Second, a choice has to be made as to what sort of summary statistics should be accepted as the raw input. There are two obvious choices. One is to take treatment effect estimates (typically contrasts of treatment vs control, one for each trial) with associated standard errors and the other is to take treatment means (at least two for each trial depending on the number of arms). The former has the advantage that problems of trial design can be handled off-line. All that is needed are treatment effect estimates that validly take account of the relevant designs, so that, for example, the standard error for the between treatment contrast for a cross-over trial should be based on within patient errors. Likewise, the treatment effect estimate might be adjusted for additional covariates other than treatment.

As already mentioned in the introduction, the pharmaceutical industry faces an increasing demand for meta-analyses. Up to now, such analyses were performed by writing software code (usually in SAS) on demand. Operationally, this resulted in multiplication of programming jobs (e.g. production of forest plots) replicated independently by various teams. Methodologically, it meant that different teams were using different analyses strategies. For these reasons, some standardization of routine meta-analyses was desired. As the pharmaceutical industry has limited experience with meta-analysis, but a lot of in-house expertise on the closely related field of randomised controlled clinical trials, it was felt that a suite of macros should be built based on familiar SAS procedures. Furthermore, the development steps themselves, i.e. the selection of methods to be included in (or

ACCEPTED MANUSCRIPT

excluded from) the standardized meta-analysis software package, as well as training on its use, including interpretation of analysis results and graphs, were also considered important benefits of the collaboration.

Indirect analyses, for example, so-called 'network analyses' might sometimes be easier to implement with summary statistics by trial and treatment group. However, if some trials included in the network analysis are more complex (and, as mentioned, incorporate covariate adjustments and/or cross-over designs), then a sensible estimate of a common treatment effect requires more information than just treatment means and their variances. Thus, it was decided to develop software to handle inputs in the form of treatment contrasts and leave as a possibly more ambitious objective for the future the construction of a program to carry out network analyses. What is described here, therefore, is a suite of SAS® programs to carry out various forms of fixed and random effects meta-analysis and also to produce various supplementary checks.

3. Development and Validation

Writing of the macros themselves was largely the work of JW with some input from SS but involved as a preliminary step a thorough evaluation of some existing packages and their functionality. Packages reviewed were CMA® (Comprehensive Meta-Analysis), MIX® (Meta-analysis with Interactive eXplanations), RevMan®, (the Cochrane Collaboration software), R® (using the meta and rmeta modules), GenStat® version 10 (the Meta procedure), MetaWin® and Stata®. These packages were studied to see what features would be desirable but were also used as part of an initial informal validation of algorithms by providing benchmarks against which results could be checked.

The Novartis team went through all formal validation steps as required of the pharmaceutical industry to be fully compliant with the US Food & Drug administration's 21 CFR Part 11 [see reference: FDA 21 CFR Part 11: Code of Federal Regulations; Electronic Records; Electronic Signatures; 20March1997] and to ensure that the macros run reliably in the Novartis programming environment. Validated macros certainly best support an efficient operational process and allow future macro updates (removing bugs, implementing macro improvements). As part of the validation plan, the Novartis team set up, among other things, user requirement specifications, functional specifications and test cases. The latter consisted of a number of sample applications of the macros to trial data matching the user requirements. The results produced by the macros were compared with published results or calculations performed in other software packages (mostly R).
All validation activities, including the test results were documented for potential health authority inspection.

Of course, closely allied to the process of validation, is that of establishing that the macros were fit for purpose from the point of view of the end-user. This involved, in addition to communication by email and teleconference, presentations by SS and JW in Switzerland as the macros were being developed and then later SS in giving courses in their use in both Switzerland and the USA.

ACCEPTED MANUSCRIPT

4. The macros

The various modules are being made freely available in 'pre-production form' (that is to say in a form prior to their final validated adoption by Novartis). Please refer to the disclaimer outlined within the Introduction regarding user liabilities.
In describing the modules it will be useful to remember that both basic fixed and random estimators are of the general form.

$$\hat{\theta}_{meta} = \sum_{i=1}^{k} w_i \hat{\theta}_i, \ \sum_{i=1}^{k} w_i = 1 \tag{1}$$

with variance given by

$$V\left(\hat{\theta}_{meta}\right) = \sum_{i=1}^{k} w_i^2 V\left(\hat{\theta}_i\right), \tag{2}$$

where $V\left(\hat{\theta}_i\right)$ is the variance of the estimate from the ith trial but the w_i depend on the method chosen.

The modules that were created were as follows.

1. MAINVERSE – To carry out classic fixed effects meta-analysis using inverse weighting by variances of treatment contrasts. This approach is the classic meta-analytic fixed effect approach to combining results of trials when these themselves are summarised in terms of a treatment contrast, presumed approximately Normally distributed, and an estimated standard error. Thus if the estimated effect in trial $i, i = 1, \ldots, k$ is $\hat{\theta}_i$ and its estimated standard error is $\hat{\phi}_i$ the overall estimate is of the general form given in (1) but with $V\left(\hat{\theta}_i\right) = \phi_i^2$ and

$$w_i = \frac{1/\hat{\phi}_i^2}{\sum_{i=1}^{k} 1/\hat{\phi}_i^2}. \tag{3}$$

   It is described in texts such as, for example(Petitti, 2000; Whitehead, 2002).

2. MARANDOM – To carry out random effects analysis using the approach of DerSimonian and Laird (DSL)(DerSimonian and Laird, 1986) and also using that of Hardy and Thompson (HT)(Hardy and Thompson, 1996). In both of these cases, the within-trial estimate of the variance of the treatment contrast is augmented using a random effect variance estimate so that the total variance for a trial is $V\left(\hat{\theta}_i\right) = \phi_i^2 + \tau^2$ and weights in (1) must be of the form

ACCEPTED MANUSCRIPT

$$w_i = \frac{1/\left(\hat{\phi}_i^2 + \tau^2\right)}{\sum_{i=1}^{k} 1/\left(\hat{\phi}_i^2 + \tau^2\right)} \tag{4}$$

They differ in that the DSL approach provides a moment-based estimate of τ^2, whereas the HT approach is a maximum likelihood approach requiring the iterative solution of two simultaneous equations. MARANDOM also produces shrunk trial effects and provides the user with confidence intervals of the Wald and profile likelihood form, if the HT approach is selected.

3. MABINARY – To carry out various approaches to analysing binary data including classic Mantel-Haenszel analysis(Mantel and Haenszel, 1959), but also, for example, the analysis of Normal-binomial mixtures using PROC NLMIXED®. In the latter case, the number of events in arm $j, j = 1, 2$, of trial i are supposed to be distributed as a binomial so that the number of events

$$Y_{ij} \sim Bin\left(\pi_{ij}, n_{ij}\right) \tag{5}$$

where π_{ij} is the probability of an event and n_{ij} is the number of subjects in the given arm and trial. It is supposed however that if $\eta_{ij} = \log\left[\pi_{ij}/\left(1-\pi_{ij}\right)\right]$ is the logit of the event probability. This in turn may be modelled as

$$\eta_{ij} = \mu_i + \theta_i z_j \tag{6}$$

where

$$\theta_i \sim N\left(\theta, \sigma^2\right) \tag{7}$$

is a random treatment effect, assumed Normally distributed and $z_1 = 0, z_2 = 1$ is a treatment indicator.

4. MAFOREST – To produce so-called forest plots(Lewis and Ellis, 1982; Lewis and Clarke, 2001) whereby individual trials are represented by horizontal lines joining lower and upper confidence limits with a plotting symbol for the estimate and a final line and symbol for the overall meta-analytic summary. A different vertical plotting position is chosen for each trial.

5. MAFUNNEL – To produce funnel plots(Begg, 1985; Egger et al., 1997; Vandenbroucke, 1988). These plot the treatment estimate on the horizontal axis and the reciprocal of the standard error of the treatment estimate on the vertical axis. If the actual treatment effects observed are independent of the estimated standard errors, then they should cluster around the average estimated treatment effect, the degree of scatter being higher the higher the standard error. Hence the points should approximately fill the shape of a cone with its funnel pointing upwards. If a part of this cone is not filled it may

ACCEPTED MANUSCRIPT

imply some possible publication bias whereby, for example, only positive studies appear where the standard error is high.

6. MAGALBRAITH – To produce radial or Galbraith plots(Galbraith, 1988a, b, 1994). These plot the Z-scores, that is to say the ratio of estimated treatment effect to standard error, against the reciprocal of the standard error, where the latter is calculated as for a fixed effects analysis. Thus we have a plot of

$$Z_i = \frac{\theta_i}{\phi_i}, \textit{versus } P_i = \frac{1}{\phi_i}. \tag{8}$$

It then follows that the ordinary least squares estimate, $\hat{\beta}$, of the slope, through the origin, β of the regression of Z on P is the fixed effects meta-analytic estimate. To see this note that

$$\hat{\beta} = \frac{\sum_{i=1}^{k} Z_i P_i}{\sum_{i=1}^{k} P_i^2} = \frac{\sum_{i=1}^{k}\left(\frac{\hat{\theta}_i}{\phi_i}\frac{1}{\phi_i}\right)}{\sum_{i=1}^{k}\left(\frac{1}{\phi_i}\right)^2} = \frac{\sum_{i=1}^{k}\left(w_i\hat{\theta}_i\right)}{\sum_{i=1}^{k} w_i} = \hat{\theta}_{fixed}, \tag{9}$$

since the form of the w_i in (9) is as given in (3). The plot is useful for assessing heterogeneity.

7. MAQQ – To produce QQ plots of estimated treatment effects by trial. Here each deviation from the overall estimate is expressed as a standardised difference using either the fixed effect standard error, ϕ_i or the random effect standard error $\sqrt{\phi_i^2 + \tau^2}$. This is useful for examining whether it is plausible that the estimates may be regarded as coming from a single Normal distribution. Of course, where the within-trial standard errors differ, this cannot be the case. Nevertheless the resulting mixture may to a greater or lesser extent be reasonably approximated by a Normal distribution.

8. MAPETERLEE – To implement Lee's checks(Lee, 1999). Lee proposed a number of useful plausibility checks of summary statistics from studies with binary outcomes. For example, since an approximate formula for the variance of a log-odds ratio is

$$V(\theta) = \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}, \tag{10}$$

and since for given total subject numbers N, the variance $V(\theta)$ reaches a minimum when

$$a = b = c = d = \frac{N}{4}, \tag{11}$$

substitution of (11) in (10) yields

ACCEPTED MANUSCRIPT

$$V\left(\hat{\theta}\right) \geq \frac{16}{N},\ N \geq \frac{16}{V\left(\hat{\theta}\right)} \tag{12}$$

and since $V\left(\hat{\theta}\right)$ can readily be calculated from the confidence intervals for a given study, such intervals immediately suggest a lower bound on the number of subjects. Similar bounds can be obtained for the numbers on either treatment or control or resulting in 'success' or 'failure'. This macro calculates such bounds, prints them out in a table by study and also plots them. This is useful to detect errors in the summary input data or in the publications when the original tables of counts are not available.

9. MASENSITIVITY – To examine the sensitivity of conclusions from random effects analysis to the magnitude of the random effects variance. A common problem with random effects meta-analysis is that the between trial component of variation is badly estimated, since the degrees of freedom for estimating this are $k-1$. It is often the case that the number of trials k is far fewer than the number of patients in the smallest trial. Furthermore, unlike the within-trial variance, the random effect variance cannot be estimated directly. Thus of all the variances estimated in a meta-analysis this is frequently the least well estimated. This module allows the user to try various values and see to what extent the resulting inference is modified.

The ways that the modules interconnect is described in Figure 1. For many applications it will not be necessary to use all modules. In what follows we describe an example where they could all be used.

<Figure 1 about here>

5. An example

As an example to illustrate the use of the modules consider the meta-analysis of trials in pre-eclampsia carried out by Collins et al. (Collins et al., 1985) and also reported in the book by Brown and Prescott(Brown and Prescott, 2006). These data will be referred to as pre-eclampsia data from now on.

<Figure 2 about here>

Figure 2 represents a SAS® data step inputting the pre-eclampsia data and storing them in a SAS dataset pre_eclamp.

<Figure 3 about here>

Next, we assume that the nine modules listed above have been compiled. Then the calls given in Figure 3 are made. Note that the order in which these are made is important. In particular the graphical macros require that suitable datasets have been prepared first. Because the original counts are available, the analysis does not need to proceed using summary statistics. There is no point in this case, therefore in performing Peter Lee's checks. The call is included in the code purely for illustrative

ACCEPTED MANUSCRIPT

purposes. Also, it could be argued that the fixed effects and random effects analyses using summary statistics, which are produced by calls to mainverse and marandom respectively are not needed since an analysis on the binary scale has been carried out. That is to say a fixed effects version has been carried out using Mantel-Haenszel type approaches (illustrated in Figure 4) and a random effects approach using a normal-binomial mixture. Nevertheless, there may be some interest in comparing this to the other common approach using summary contrasts and in any case these calls are included for illustrative purposes.

<Figure 4 about here>

Table 1 gives a summary of various fixed and random effects estimates, standard errors and confidence intervals for the overall treatment effects on the log-odds scale produced by the macros. Two fixed effect and four random effect approaches are given as well as the result of applying a logistic regression analysis using proc logistic. This particular option is easily programmable separately and is not included in the macros, although sample code is included on the website in the program used to run the macros. This thus constitutes a third fixed effects approach. Two of the random effect approaches differ only in the way that confidence intervals are produced. These are Hardy-Thompson(Hardy and Thompson, 1996) likelihood approaches differing only as to whether confidences limits are produced as Wald limits or profile likelihood limits.

It is noticeable that the three fixed effect approaches produce very similar results, with estimates varying from -0.398 to -0.410 with standard errors of either 0.088 or 0.089. The random effect estimates are rather different from the fixed effects ones but similar to each other with values ranging from -0.526 to -0.517 and standard errors very much higher and ranging from 0.186 to 0.206. The Galbraith and funnel plots given in Figures 5 and 6 provided alternative ways of illustrating the heterogeneity in this example.

<Table 1 about here>

<Figure 5 about here>

The very much higher standard errors for the random effect approaches reflect the heterogeneity in the data. The Q statistic is 27.3 on 8 degrees of freedom. The random effects variance estimates for the DSL and HT methods are very similar at 0.23 and 0.24 with the values for the normal binomial mixture somewhat lower at 0.18. However these values are large compared to the typical within trial variance, the median of which is 0.12 and the lowest of which is 0.015. However, as pointed out by Senn(Senn, 2007), this alone is not sufficient to make the fixed effects and random effects estimates differ. It is also necessary that there is a difference between the unweighted mean of the effect estimates from the original trials and the fixed effects estimate as well as a correlation between the variances and the effect estimates from the trials. The unweighted mean is -0.467 here and the correlation between variances and estimates is 0.48.

<Figure 6 about here>

ACCEPTED MANUSCRIPT

The difference between the fixed effects and random effects solutions is reflected in the sensitivity plot shown in Figure 7. This plots effect estimates and confidence intervals against the random effects variance. Where the variance is set to zero, a fixed effects solution is produced. Thus, the middle curve starts at -0.4. However at the value of the random effects variance of 0.23, the value reaches -0.5.

<Figure 7 about here>

Some would hold strongly to the view that heterogeneity is a reason for preferring the random effects approach. However, our view is that the difference is one of purpose, with the fixed effects approach being a means of examining whether the treatment can have an effect and if so what it was in the patients actually studied and the random effects approach more relevant for predicting what might happen in further trials(Senn, 2000). Of course, if heterogeneity is low, both approaches give very similar answers and therefore conversely high heterogeneity is an indication, at least, that a fixed effects analysis is unlikely to give an adequate answer to the sort of question that random effects analyses are designed to address.

6. Discussion

Obviously there are pros and cons to having a set of macros to cover aspects of meta-analysis.

The pros include the following. First, although there is a considerable investment of time in creating the macros, in the long run, if similar analyses are often performed, there is a return in terms of saved time and efficiency. Second, there may be some advantages in terms of reliability since code can be ‘validated’ once and for all and will then deliver correct results. Whatever ones cynicism about this process, (anybody who uses software regularly will have made discoveries of bugs in ‘validated’ software) one can at least hope that repeated use of the same algorithm will lead to useful feedback, which can be used either to increase confidence in the approach or information as to how it needs to be corrected. In that spirit, of course, we are interested to get feedback from users of the macros. Third, by making it easy to perform a number of different analyses they may encourage wider exploration of datasets. Fourth, they may make it possible for those who are familiar with basic application areas but less expert on coding to concentrate on more important aspects of the problem.

Inevitably there are some drawbacks. The use of a suite of macros may encourage routine application of what is easily available at the expense of thinking about what is appropriate. There is some evidence that the spread of software for meta-analysis has caused enthusiasm to exceed expertise in some cases(Senn, 2009). Second, the fact that certain standard analyses are available may discourage would-be analysts from making the effort to explore and code alternatives. We are aware that this can be a problem and have done what we can to make exploration a possibility, providing, for example routines for qq plots, sensitivity analysis and carrying out Lee’s checks(Lee, 1999). Nevertheless, one can only go so far in this and anybody wishing to carry out a

ACCEPTED MANUSCRIPT

meta-analysis in SAS should regard these macros as a possible (we hope useful) tool but by no means as the end of the story.

We hope that others find these macros useful and look forward to receiving feedback on their use.

Acknowledgement

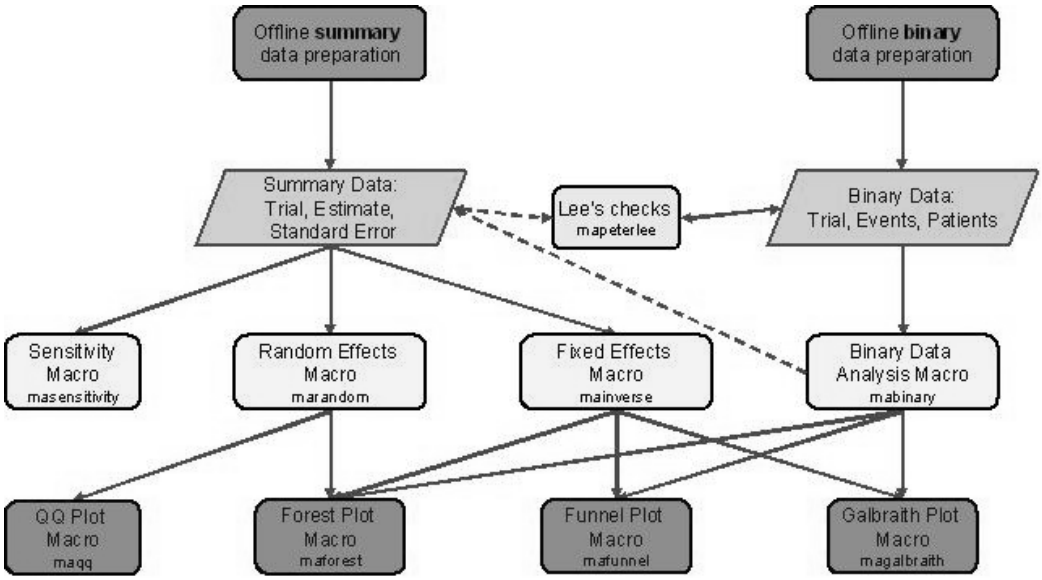
We are grateful to a referee and the editor for helpful comments on an earlier draft.

References

Bohlius, J., Schmidlin, K., Brillant, C., Schwarzer, G., Trelle, S., Seidenfeld, J., Zwahlen, M., Clarke, M., Weingart, O., Kluge, S., Piper, M., Rades, D., Steensma, D.P., Djulbegovic, B., Fey, M.F., Ray-Coquard, I., Machtay, M., Moebus, V., Thomas, G., Untch, M., Schumacher, M., Egger, M., Engert, A., 2009. Recombinant human erythropoiesis-stimulating agents and mortality in patients with cancer: a meta-analysis of randomised trials. Lancet 373, 1532-1542.
Brown, H., Prescott, R., 2006. Applied Mixed Models in Medicine, Second ed. Wiley, Chichester.
Collins, R., Yusuf, S., Peto, R., 1985. Overview of randomized trials of diuretics in pregnancy. BMJ 290, 17-23.
DerSimonian, R., Laird, N., 1986. Meta-analysis in clinical trials. Controlled clinical trials 7, 177-188.
Hardy, R.J., Thompson, S.G., 1996. A likelihood approach to meta-analysis with random effects. Statistics in Medicine 15, 619-629.
Lee, P.N., 1999. Simple methods for checking for possible errors in reported odds ratios, relative risks and confidence intervals. Statistics in Medicine 18, 1973-1981.
Mantel, N., Haenszel, W., 1959. Statistical apsect of the analysis of data from retrospective studies of disease. Journal of the National Cancer Institute 22, 719-748.
Normand, S.L., 1999. Meta-analysis: formulating, evaluating, combining, and reporting. Statistics in Medicine 18, 321-359.
Petitti, D.B., 2000. Meta-Analysis, Decision Analysis and Cost-Effectiveness Analysis, Second ed. Oxford University Press, Oxford.
Senn, S., 2010. Hans van Houwelingen and the art of summing up. Biometrical Journal 52, 85-94.
Senn, S.J., 2000. The many modes of meta. Drug Information Journal 34, 535-549.
Senn, S.J., 2007. Trying to be precise about vagueness. Statistics in Medicine 26, 1417-1430.
Senn, S.J., 2009. Overstating the evidence: double counting in meta-analysis and related problems. BMC Med Res Methodol 9, 10.
van Houwelingen, H.C., Arends, L.R., Stijnen, T., 2002. Advanced methods in meta-analysis: multivariate approach and meta-regression. Statistics in Medicine 21, 589-624.
Whitehead, A., 2002. Meta-analysis of Controlled Clinical Trials. Wiley, Chichester.

ACCEPTED MANUSCRIPT

Figure



ACCEPTED MA

ACCEPTED MANUSCRIPT

Figure

```
/* Example data set as found on page 204 of        */
/* Brown, H. and R. Prescott (2006). Applied Mixed */
/* Models in Medicine. Chichester, Wiley.          */

/* Input data and save in dataset pre_eclamp */

data pre_eclamp;
input study $ treat_event Nt cont_event Nc;
datalines;

Weseley        14  131  14   136
Flowers        21  385  17   134
Menzies        14   57  24    48
Fallis          6   38  18    40
Cuadros        12 1011  35   760
Landesma      138 1370  175 1336
Krans          15  506   20  524
Tervila         6  108   2   103
Campbell       65  153  40   102;
```

ACCEPTED MANUSCRIPT

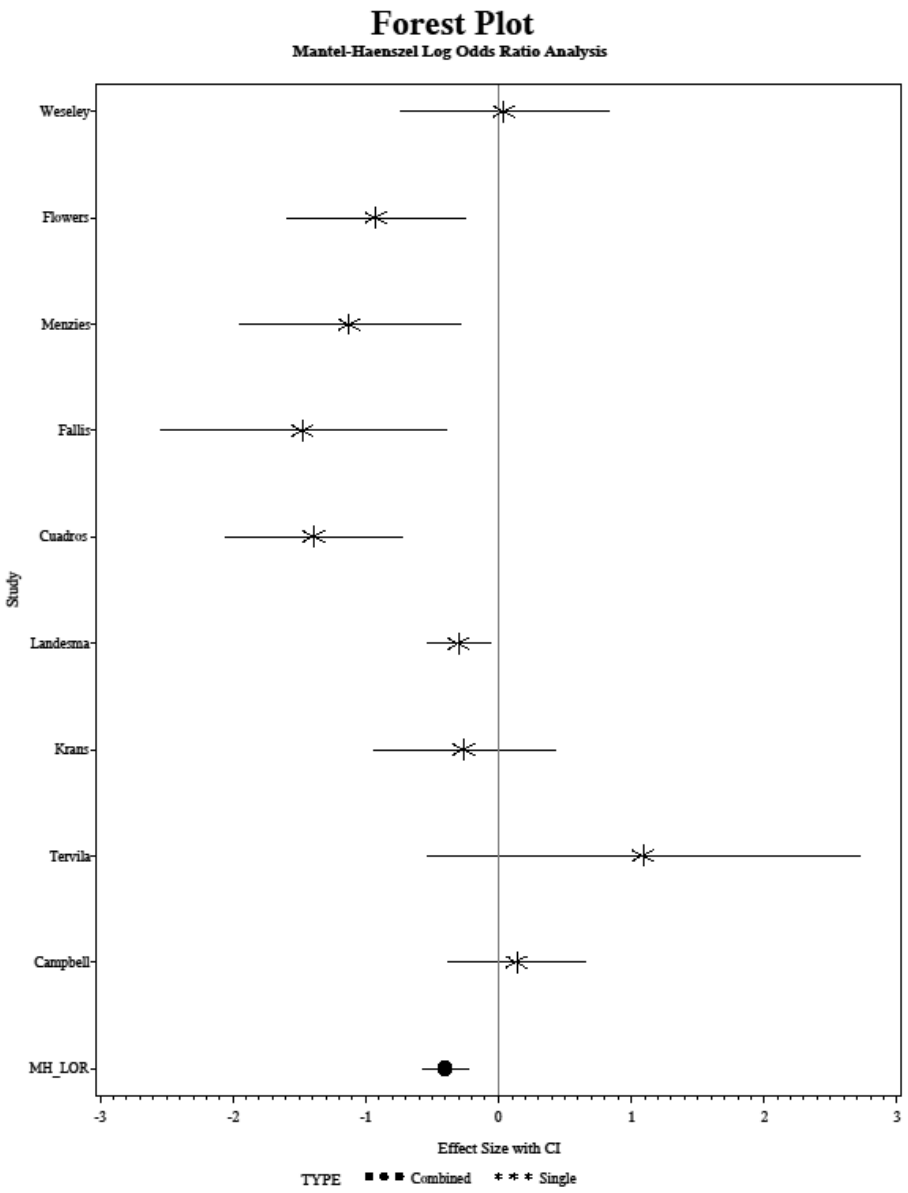
ACCEPTED MANUSCRIPT

Figure

```
/* NB remember to run all the nine meta-analysis   */
/* macros before invoking them in the code below.  */

/* Run binary macro specifying dataset pre-eclamp. */
%mabinary(dataset=pre_eclamp,method=MH_lor)
/* MH log-odds ratio analysis chosen. */

/* NB macro above will store results in datasets mh_lor and */
/* mh_lor2. Optional display of created datasets follows.   */
proc print data=mh_lor;
run;
proc print data=mh_lor2;
run;

/* maforest, mafunnel, magalbraith below use mh_lor */
/* created by mabinary as input. */
%maforest(dataset=mh_lor,transform=NO)
%mafunnel(dataset=mh_lor)
%magalbraith(dataset=mh_lor)

/* The following four calls refer to the */
/* dataset mh_lor2 created by mabinary.  */
/* Perform inverse weighting meta-analysis using   */
/* the output from the MH log-odds ratio analysis. */
%mainverse(dataset=mh_lor2,sampsize=YES) /*Classic fixed effects
analysis*/
%mapeterlee(dataset=mh_lor2) /*Peter Lee's checks*/
%masensitivity(dataset=mh_lor2,low=0.01,upp=0.29,step=0.001)
/*Sensitivity plot*/
%marandom(dataset=mh_lor2,method=BOTH,iter=100,sampsize=YES)/*Random
effects analysis*/

/* marandom created datsets qq_dsl and qq_ht. */
/* Illustrate use of maqq on DerSimonian and Laird analysis. */
%maqq(dataset=qq_dsl)

/* Now illustrate an alternative binary analysis. */
%mabinary(dataset=pre_eclamp,method=NLMIXED) /*Normal
binomial mixture chosen*/
```

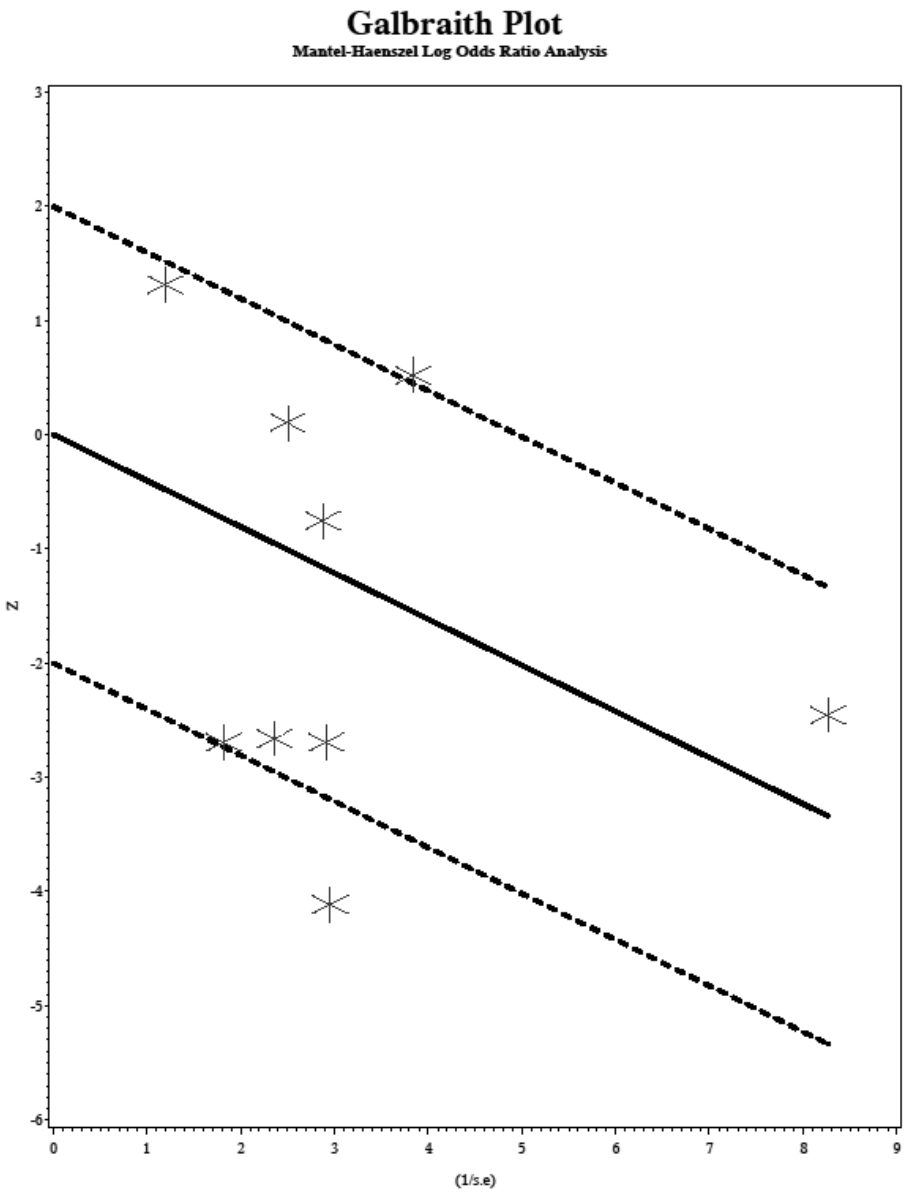
ACCEPTED MANUSCRIPT

Figure



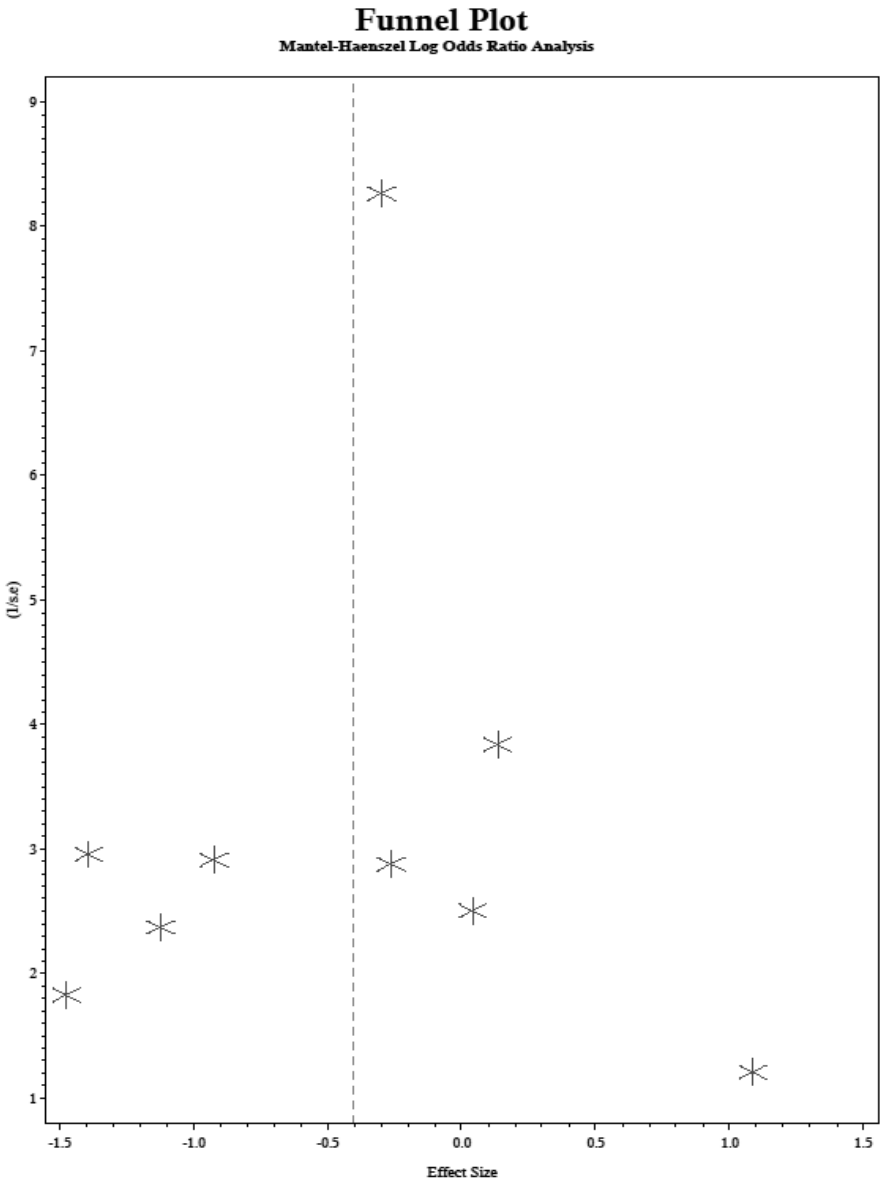
ACCEPTED MANUSCRIPT

Figure



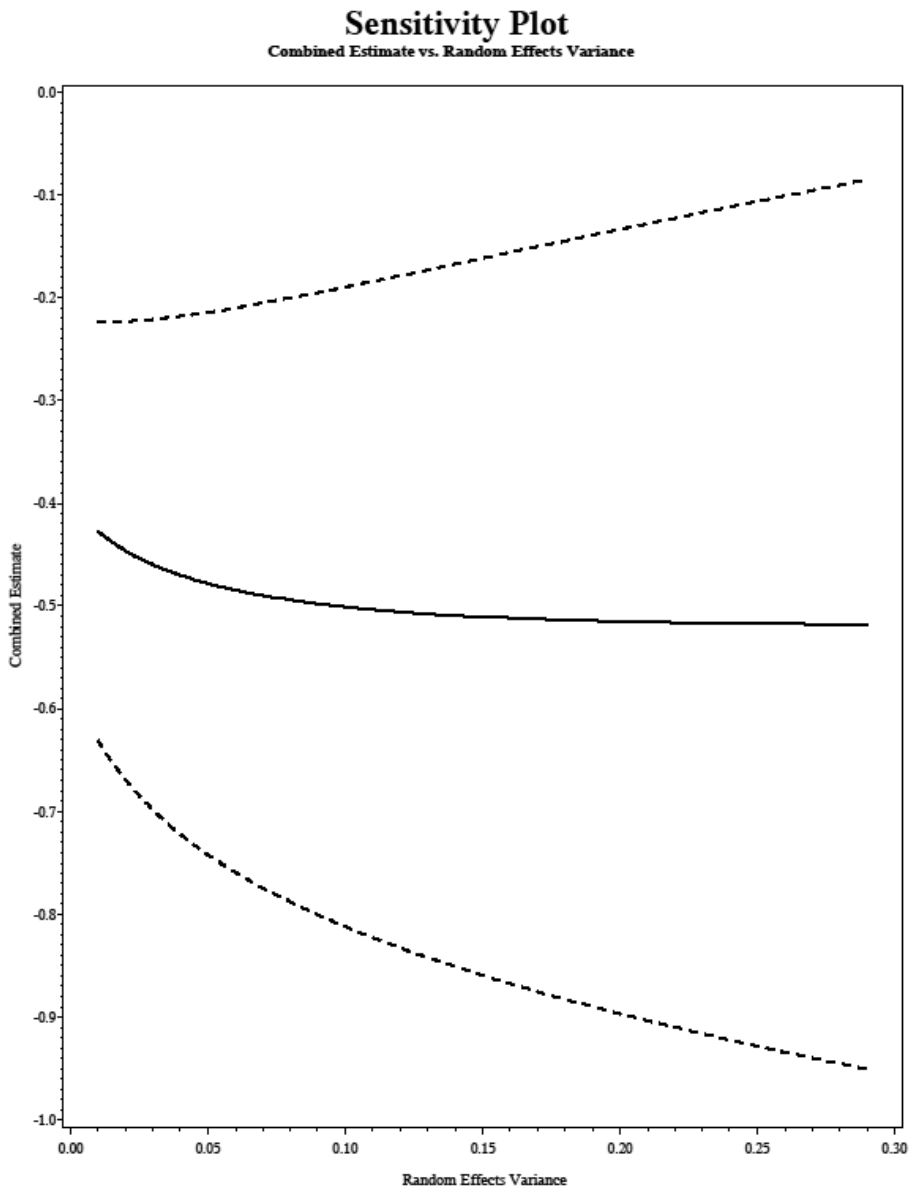
ACCEPTED MANUSCRIPT

Figure



ACCEPTED MANUSCRIPT

Figure



ACCEPTED MANUSCRIPT

Table

Table 1 Estimates, standard errors and 95% confidence limits obtained by using the macros corresponding to various approaches plied to the pre-eclampsia data

Fixed or Random	Method	Estimate	Standard Error	Lower Confidence Limit	Upper Confidence Limit
Fixed	Mantel-Haenszel	-0.404	0.088	-0.576	-0.232
Fixed	Inverse variance weighting	-0.398	0.089	-0.573	-0.223
Fixed	Logistic regression	-0.410	0.089	-0.584	-0.237
Random	DerSimonian-Laird	-0.517	0.204	-0.916	-0.117
Random	Hardy-Thompson Wald	-0.517	0.206	-0.922	-0.112
Random	Hardy-Thompson Profile			-0.983	-0.049
Random	Normal binomial mixture	-0.526	0.186	-0.956	-0.097

ACCEPTED MANUSCRIPT