

A gentle introduction to meta-analysis

Angelo Tinazzi, SENDO-Tech S.r.l., Milan, Italy
Jayne Tierney, UK Medical Research Council, London, UK

ABSTRACT

Meta-analysis (MA) to combine the results of related randomised controlled trials, has become a common and widely accepted tool in the evaluation of health care interventions. Regulatory agencies, such as the FDA, have started considering it as key supporting evidence.

MA can be based on aggregate data extracted from the reports of published trials only, aggregate data collected from all trials or centrally collected updated individual patient data (IPD) from all randomized trials.

This paper is designed to introduce the main aspects of planning and conducting a MA; the main statistical measures will be introduced for the different kind of outcomes (dichotomous, continuous and time-to-event). An overview of the main features of the statistical software supporting MA will be also reported.

INTRODUCTION

The volume of medical literature that health care professionals, researchers and policy makers need to consider, is constantly expanding, . This makes it difficult for any individual to be aware of the state of current knowledge about a particular health care intervention. For this reason, reviews of medical literature have become essential tools for keeping abreast of the new evidence that is accumulating in his or her field of interest. Moreover, reviews also identify areas where the available evidence is insufficient and further studies are required¹.

Wider recognition of the key role of reviews in synthesizing and disseminating the results of research has prompted people to consider the validity of reviews. In the 1970s and early 1980s, psychologists and social scientists drew attention to the systematic steps needed to minimize bias and random errors in reviews of research^{2,3,4,5,6}, but it was not until the late 1980s that people drew attention to the poor scientific quality of healthcare review articles^{7,8,9}.

SYSTEMATIC REVIEWS AND META-ANALYSIS

Systematic reviews use explicit, objective and prospectively-defined methods to collect, critically evaluate and synthesize studies, making them less biased and more reproducible than traditional reviews. A systematic review may, or may not, include a meta-analysis (MA): a statistical pooling of the results from individual studies to obtain a single overall estimate of treatment effect¹⁰. Meta-analysis can provide more precise estimates of the effects of an intervention than those derived from the individual studies included in a review^{11,12,13,14,15,16} and allows decisions that are based on the totality of the available evidence. The distinction between systematic review and meta-analysis is important because it is always appropriate and desirable to systematically review a body of data, but it may sometimes be inappropriate, or even misleading, to statistically pool results from separate studies. It should be noted that a number of terms have been used to describe the process of systematically reviewing and integrating research evidence, including "systematic review", "meta-analysis", "research synthesis", "overview" and "pooling".

To examine the relative effect of different interventions, ideally systematic reviews and meta-analysis should be based on randomised controlled trials. However, systematic reviews and meta-analyses of other study types such as prospective cohorts or diagnostic studies are possible and methods have been developed.

Recognition of the need for systematic reviews (with or without meta-analyses) has grown rapidly. Largely as through the efforts of organizations such as the Cochrane Collaboration, that systematically reviews across all health care areas, recognition of the need for systematic reviews has grown. This is reflected by the number of systematic reviews and empirical studies of the review methodology that have been published e.g. in the medical literature, Cochrane Database of Systematic Reviews and the Cochrane Methodology Register..

PLANNING A SYSTEMATIC REVIEW (SR)

Like a clinical trial, a SR protocol is needed before the SR can be performed. A good protocol should outline at least the following section:

- Background
- The explicit question or questions to be addressed
- Inclusion and exclusion criteria for studies
- Methods for identifying studies including search strategies and time period covered;
- Methods for quality assessment of studies, such as the randomization methods used;
- Aggregate data to be extracted or data or individual patient data to be collected;
- All planned analyses;
- Other sections as required

DEFINING THE QUESTION

This involves deciding about the intervention or interventions of interest and the participant population. The question may then be broad e.g. the effect of chemotherapy on patients with lung cancer, which would examine all types of chemotherapy in a types of patients with lung cancer. Instead you may want to ask a more or specific question e.g. the effect of cisplatin-based chemotherapy on patients with advanced non-small cell lung cancer.

INCLUSION CRITERIA

The inclusion criteria need to describe explicitly the type of study design to be included, the types of intervention and the type of patient, and possibly the time period, in order to address the question of interest. For a question concerning the effect of chemotherapy on patients with advanced non-small cell lung cancer inclusion criteria might be:

- Randomised controlled trials since 1960
- Comparison of chemotherapy with no chemotherapy
- Patients with advanced non-small-cell lung cancer

SEARCHING STRATEGIES

In order to be comprehensive and not overly selective, searches should ideally be based on a number of study sources (i.e. MEDLINE, EMBASE Cochrane Central Register of Controlled Trials, the scientific literature index portal such PubMed, Internet engine search such as google, etc). Optimal strategies for identifying trials in MEDLINE and EMBASE are available. To identify trials specific to the question, it is useful to use a mixture of subject headings and text strings. However, such searches may miss some important studies. Therefore, these should be supplemented by other searches, of the reference lists of identified study reports, relevant conference proceedings, and sources of unpublished studies (e.g. trial registers).

DATA EXTRACTION/COLLECTION

EXTRACTING DATA FROM THE ARTICLES/REPORTS IDENTIFIED

Most systematic reviews and meta-analysis rely on extracting aggregate trial data from trial reports or presentations. Data on the participants and outcomes of interest should be extracted on to a standard form and ideally it should be done independently by two individuals, who then resolve any discrepancies together. The data that can be extracted for descriptive purposes or meta-analysis will depend on the quality of reporting of in trial publications, which could be incomplete, out of date and potentially biased.

INDIVIDUAL PATIENT DATA (IPD) VS SUMMARY PATIENT DATA (SPD) META-ANALYSIS

When a meta-analysis is based on centrally collected and updated individual patient data (IPD) obtained from the trial investigators, summary statistics are potentially derived from complete and up-to-date data on all randomized patients. The collection and validation of IPD can take longer and is more resource-intensive, but the quality of the data is improved and the range of possible analyses is greater.

STATISTICAL METHODS

Most commonly, meta-analysis follows a 2-stage approach. In the first stage, summary statistics are generated for each trial and in the second stage, these statistics are summarized (pooled) in a final estimate of the treatment effect as a weighted average of the treatment effects estimated in the individual trials using the formula reported below:

$$meta - analysis\ estimate = \frac{\sum (estimate\ of\ effect\ for\ trial * weight)}{\sum (weight)}$$

ESTIMATING SINGLE TRIAL RESULTS

DICHOTOMOUS OUTCOMES

For dichotomous outcomes, the results of a trial can be presented in a 2X2 table, which indicates those participants do and do not experience an event on the research intervention and control.

Study <i>i</i>	Event	No Event	Group size
Intervention	a_i	b_i	n_{1i}
Control	c_i	d_i	n_{2i}
			$N_i (n_{1i} + n_{2i})$

Together with the number of participants in each group a number of relative effect measures can be calculated: the relative risk (RR), odds ratio (OR) and risk difference (RD). For example, based on the above table, the three measures are calculated as follows for each study (i):

$$RR_i = \frac{a_i / n_{1i}}{c_i / n_{2i}} \quad OR_i = \frac{a_i d_i}{b_i c_i} \quad RD_i = \frac{a_i / n_{1i}}{b_i / n_{2i}}$$

PhUSE 2007

An OR or a RR greater than 1 indicates that the condition or event is more likely in treatment A (if treatment A is used as a reference in the above calculation), while an OR or a RR less than 1 indicates that the condition or event is less likely in treatment A. The main difference is that OR, since it varies from 0 to ∞ , it approximates the normal distribution, while RR varies from 0 to 1; moreover the OR can give a better interpretation in the clinical context, but it may be not appropriate in situations with few or many events.

CONTINUOUS OUTCOMES

For the continuous outcomes results of a trial can be expressed in terms of the mean and associated standard deviation. Two continuous effects measures can be calculated: the difference in means, which is termed the mean difference (MD), can be used when all of the trials measure an outcome on the same scale. Where trials each measure the same outcome on a different scale of measurement, the difference in means relative to the variability observed in the trial can be used instead. This is termed the standardized mean difference (SMD).

Study i	Mean Response	Standard Deviation	Group size
Intervention	m_{1i}	SD_{1i}	n_{1i}
Control	m_{2i}	SD_{2i}	n_{2i}
			$N_i(n_{1i} + n_{2i})$

For example, based on the above table, the three measures are calculated as follows for each study (i):

$$MD_i = m_{1i} - m_{2i} \quad SMD_i = \frac{m_{1i} - m_{2i}}{s_i} \left(1 - \frac{3}{4N_i - 9} \right)$$

TIME-TO-EVENT OUTCOMES

For time-to-event outcomes, the results of a trial can be presented as a series of 2x2 tables for each time an event takes place. Each table indicates those participants do and do not experience an event on the research intervention and control. Each time (t) an event occurs, the expected events are calculated both on treatment (E_{1ti})

$$E_{1ti} = \frac{n_{1ti}(a_{ti} + c_{ti})}{N_{ti}}$$

and control (E_{2ti})

$$E_{2ti} = \frac{n_{2ti}(a_{ti} + c_{ti})}{N_{ti}}$$

The Hazard ratio (HR) for an individual trial can be then derived as follows

$$HR_i = \exp \left(\frac{\sum O_i - E_i}{\sum V_i} \right)$$

where O_i is the observed number of events in the research intervention and E_i is the expected number of events in the research intervention under the null hypothesis of equality among treatment intervention and control. V_i , the hyper-geometric variance, is calculated as follows

$$V_i = \frac{n_{1i} + n_{2i}(a_{ti} + c_{ti})(b_{ti} + d_{ti})}{N_{ti}^2(N_{ti} - 1)}$$

An for OR and RR, an HR greater than 1 indicates that the condition or event is more likely in treatment A (if treatment A is used as a reference in the above calculation), while an HR less than 1 indicates that the condition or event is less likely in treatment A.

COMBINING SINGLE STUDIES RESULTS: FIXED VS RANDOM EFFECTS MODEL

In combining single study results, two models can be approached: the fixed effects and the random effects model. Controversy still exists as the best choice of model.

FIXED EFFECTS MODEL

In this model, it is assumed that the true effect is the same or similar for each trial and that any variation is due solely to the play of chance. For example for dichotomous outcomes the Mantel-Haenszel or inverse variance approach can be used to weight the trial estimates as follows:

$$OR_{MH} = \frac{\sum w_{iMH} OR_i}{\sum w_{iMH}}$$

where w_{iMH} , the weight associated to each trial, is calculated as

$$w_{iMH} = \frac{b_i n_{1i}}{N_i}$$

RANDOM EFFECTS MODEL

When the assumption of similar true effect is not plausible for all trials, the more conservative random effect model should be taken into consideration; this model assumes that there is no one single true effect, but that there are two sources of expected variation.

EVALUATING THE HETEROGENEITY

A SR and MA can provide convincing and reliable evidence about the effects of interventions, especially when the results of the studies show effects of similar magnitude. By contrast, when the included studies have markedly different results, the conclusions are less clear.

There are a number of test statistics e.g. Cochran's Q to establish whether study results are heterogeneous, or whether the variation in results is compatible with chance alone (homogeneity). Where heterogeneity is detected it is important to explore the reasons, rather than ignore it and it may not appropriate to combine the studies.

POTENTIAL ISSUES

Publication bias and other reporting biases is a major threat to the validity of meta-analysis; if introduced, this bias may distort the evidence. Bias may also be introduced if the methodological quality of controlled trials is inadequate. It is crucial to understand the limitations of meta-analysis and the importance of exploring sources of heterogeneity and bias¹⁷.

GRAPHICAL REPRESENTATION

There are many ways of analyzing and displaying data arising from a meta-analysis and their use have different application. A brief description of the most adopted graphical representation is hereby reported.

FOREST PLOT

The "Forest plot", also known as confidence intervals plot, is the common graphical way used to reports meta-analysis results. The origin of Forest plot goes back at least to the 1970s, where Freiman et al¹⁸ displayed the results of several studies with horizontal lines showing the confidence interval for each study and a mark to show the point estimate. However, the origins of the name of this plot are obscured by history and myth.

The type of plot can be used to graphically display the results from a MA based on any of the previously described type of outcomes and therefore measures. The example in figure 1 is about a MA of a time to event outcome (e.g. Time to progression) and therefore the effect size of the treatment being evaluated is measured with the HR; the plot is composed of two distinct sections:

- on the left part the results:
 - the trial labels;
 - the number of patients and events (on treatment and control);
 - the summary statistics (in this case the observed minus expected events and the variance for individual trials, subsets and overall).
- on the right part the graphical plot:
 - the treatment effect (HR in this example) for each trial is represented by the square on the horizontal bar, the size of which is proportional to the amount of information available in the trial;
 - the inner and outer limits of the bar indicate the 95% and 99% confidence intervals respectively;
 - the centre of the black diamond gives the overall treatment effect when the results of all trials are combined. The extremes of the diamond give the 95% confidence interval.

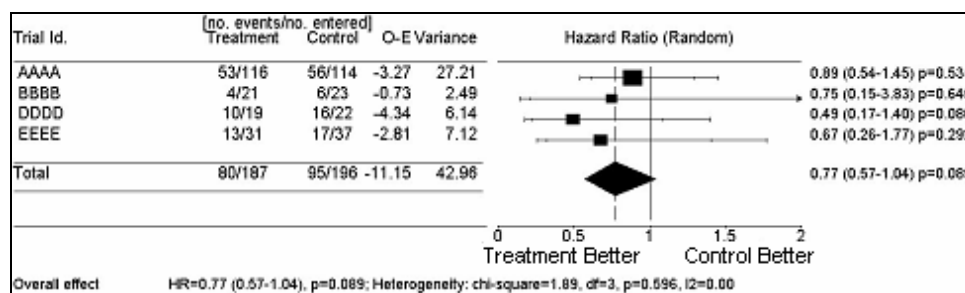


Figure 1: Forest Plot

FUNNEL PLOT

A funnel plot is a scatter plot of treatment effect against a measure of study size. It is used primarily as a visual aid to detecting bias or systematic heterogeneity.

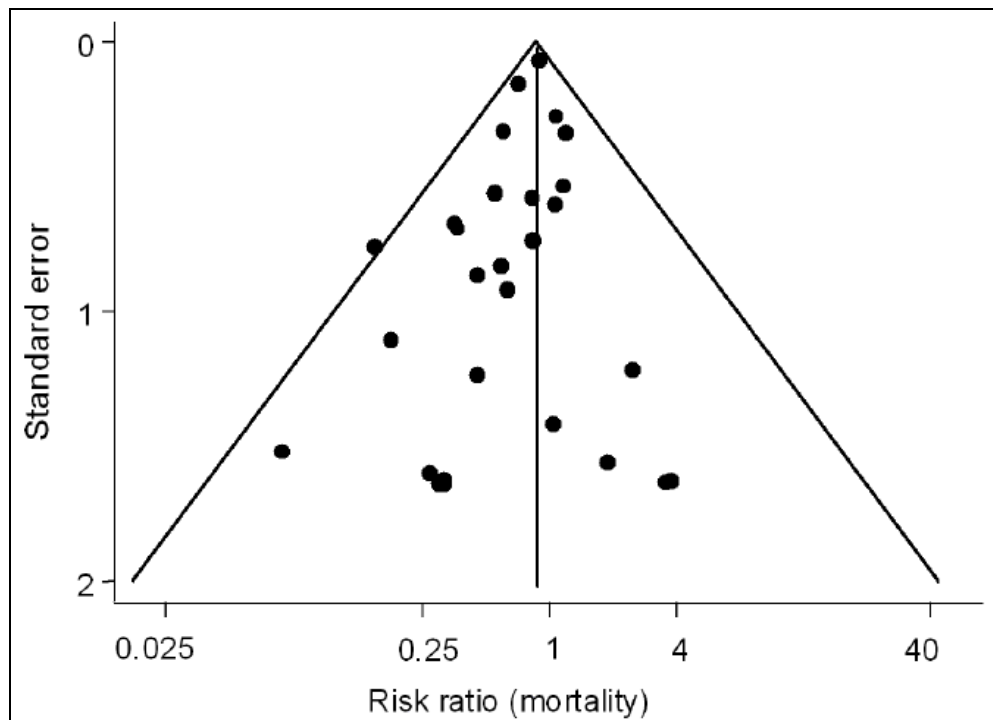


Figure 2: Funnel Plot (example from Cochrane Handbook¹)

L'ABBE' PLOT

L'Abbè is useful in analyzing dichotomous outcome. The plot shows for each study the observed event rate in the experimental group plotted against observed event rate in the control group. This plot may be used to view the range of event rates among the trials, to highlight excessive heterogeneity, and, on occasion, to indicate which treatment effect measure may be most consistent across trials.

REVIEW OF STATISTICAL SOFTWARE FOR META-ANALYSIS

A software for supporting meta-analysis techniques should take care of the following aspects:

- able to work either with aggregated or individual patient data;
- support for main effect measures presented in the previous section, applying either fixed and random effect model;
- supporting subgroup analysis;
- evaluate the heterogeneity between studies and/or subgroups (e.g. effect size in male patients vs female patients);
- graphical representation covering at least the Forest Plot and the ability to export the results (graph) to other packages such as Ms Office.

The last aspect represents the main challenge for developer as it requires the development of an uncommon graphic layout model.

Many general statistical software packages have included options for meta-analysis in their basic configuration; in addition, user-communities have written numerous meta-analysis add-on (SAS[®] macro, S+ and STATA modules). Specialized software packages for meta-analysis, are also available both with commercial and freeware license (table 2)¹⁹. Table 2 reports a summary of available specific tools for meta-analysis.

Among the existing software, Comprehensive Meta-Analysis appears the most complete and it offers a good support for continuous update (figure 4). Apart the availability of a number of useful graphical representation, it allows to enter data in more than 100 formats, that is a good option to consider when statistics to be extracted from the singles studies are not explicitly presented²⁰.

The Cochrane Collaboration has also developed a software for supporting SR (from protocol to results) and MA (calculation and plots); the software, called RevMan, is available from the Cochrane website (see table 2).

Other ad-hoc systems have been developed in the past year to solve particular issues, such as the application developed by Zamora et al²¹ and his colleagues for SR/MA of test accuracy trials (trials investigating methods / instruments for accurate diagnosis).

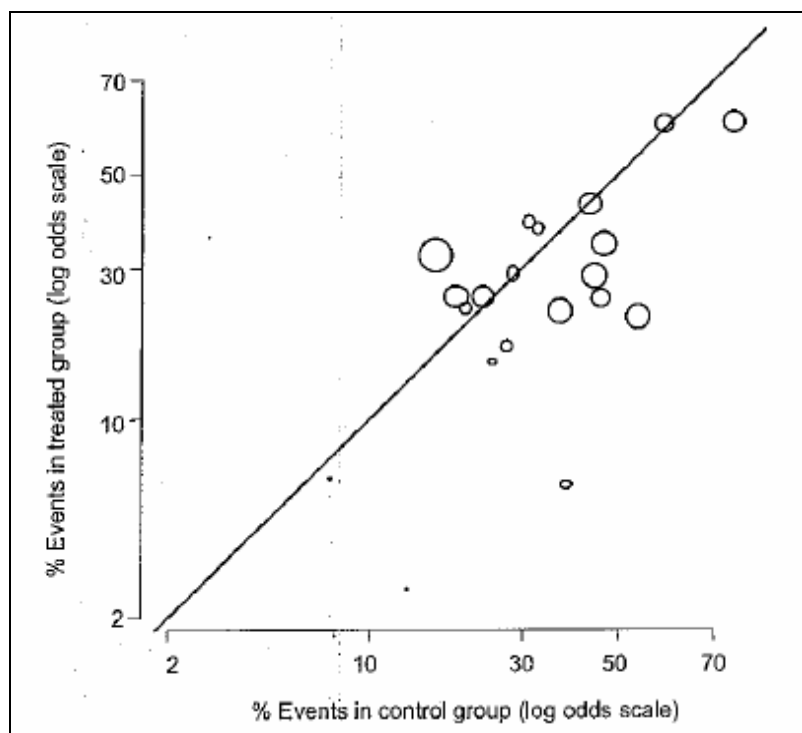


Figure 3: L'Abbè Plot (example from Cochrane Handbook¹)

SAS RESOURCES

SAS itself does not have any specific procedure or tools to support meta-analysis. However existing SAS/STAT[®] procedures such as PROC FREQ for calculating binary statistics (i.e. Odds Ratio), PROC LIFETEST for calculating time-to-event statistics (i.e. o-e and variance, or log-rank) and SAS/GRAPH[®] basic procedures combined with the annotated facility, represent a good basis for creating ad-hoc analysis program or SAS stand-alone application using the macro facility and/or other application development tools such as SAS/AF[®].

There is wide availability of SAS user macro routines available over the Internet, either supporting calculation and graphical representation^{22,23,24}. In addition the MRC Clinical Trials Units has developed a SAS based software called SCHARP²⁵ an interactive SAS based application, developed using SAS/BASE[®], SAS/STAT, SAS/GRAPH and SAS/AF modules. The application was developed to work with any of the type of outcomes previously described and since it was developed using SAS it can manage Individual Patient Data meta-analysis. The software is freely available for academic institutions following an agreement with the MRC.

Software Name	License Type
Comprehensive Meta-Analysis www.meta-analysis.com	Commercial
Meta-DiSc www.hrc.es/investigacion/metadisc_en.htm	Freeware
MIX www.mix-for-meta-analysis.info	Freeware
NCSS www.ncss.com/metaanal.html	Commercial
RevMan www.cochrane.org	Freeware
SCHARP (contact the author)	Freeware
STATA www.stata.com	Commercial
WinBUGS www.mrc-bsu.cam.ac.uk/bugs	Freeware

Table 2: Available Software for Meta-Analysis

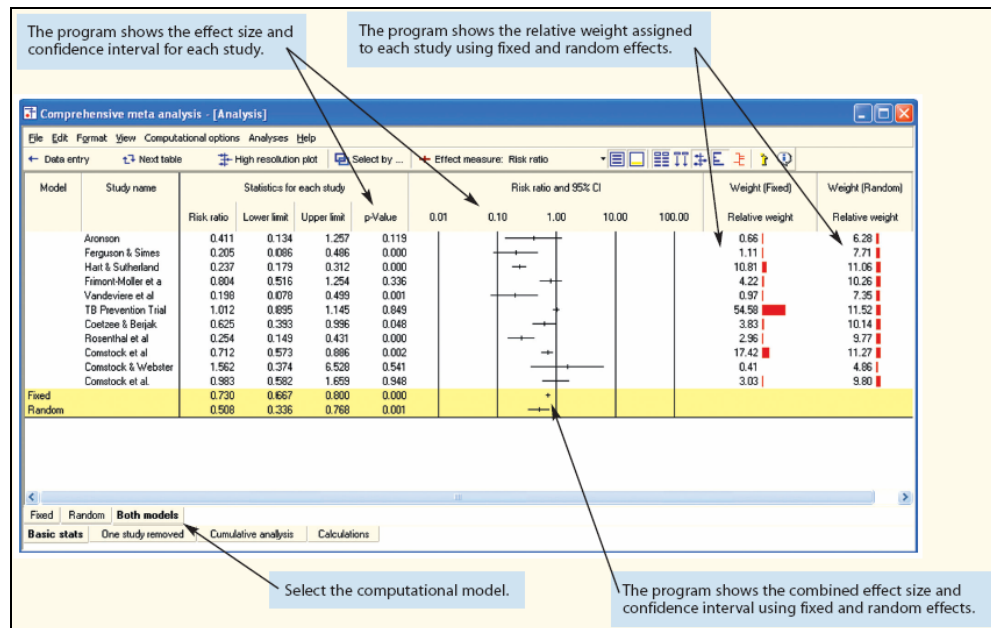


Figure 4: Comprehensive Meta-Analysis Software

CONCLUSIONS

SR and MA have become well-known methods for the synthesis of quantitative data from previously conducted research in applied health science and they can be considered the main tools for disseminating clinical evidence in the medical communities. The methodologies applied are continuously under review and The Cochrane Collaboration, named after Archie Cochrane a pioneer in health services research whose visions are at the heart of the Collaboration, is the official organization with this aim. Moreover, the role of this international collaboration is to prepare, disseminate and continuously update systematic reviews of controlled clinical trials, especially addressing many of the currently unresolved issues.

Although there are specific tools supporting MA techniques, in some circumstances they may not be appropriated. In this case, popular software such as SAS, can be 'manipulated' so that the desired calculations and/or data visualization, can be supported.

REFERENCES

1. The Cochrane Collaboration. Cochrane Handbook for Systematic Reviews of Interventions. 2006 Version 4.2.6
2. Light RJ, Smith PV. Accumulating evidence: procedures for resolving contradictions among different research studies. *Harv Educ Rev* 1971; 41:429-71.
3. Glass GV. Primary, secondary, and meta-analysis of research. *Educ Res* 1976; 5:3-8.
4. Rosenthal R. Combining results of independent studies. *Psychol Bull* 1978; 85:185-93.
5. Jackson GB. Methods for integrative reviews. *Rev Educ Res* 1980; 50:438-60
6. Cooper HM. Scientific guidelines for conducting integrative research reviews. *Rev Educ Res* 1982; 52:291-302.
7. Mulrow CD. The medical review article: state of the science. *Ann Intern Med* 1987; 106:485-8.
8. Yusuf S, Simon R, Ellenberg S (eds). Proceedings of 'Methodologic issues in overviews of randomized clinical trials'. *Stat Med* 1987; 6:217-409.
9. Oxman AD, Guyatt GH. Guidelines for reading literature reviews. *Can Med Assoc J* 1988; 138:697-703.
10. Chalmers I, The Cochrane Collaboration: preparing, maintaining and disseminating systematic reviews of the effects of health care. *Annals of the New York Academy of Science* 1993; 703:156-165.
11. Chalmers I, Altman D. Systematic Reviews. BMJ Publishing Group 1995
12. Huque MF. Experiences with meta-analysis in NDA submissions. *Proc Biopharmac Sec Am Stat Assoc* 1998; 2: 28-33.
13. Oxman AD. Meta-statistics: Help or hindrance? *ACP J Club* 1993; 118:A-13.
14. Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, Chalmers TC. Meta-analyses of randomized controlled trials. *N Engl J Med* 1987; 316:450-5.
15. L'Abbe KA, Detsky AS, O'Rourke K. Meta-analysis in clinical research. *Ann Intern Med* 1987; 107:224-33.
16. Thacker SB. Meta-analysis: a quantitative approach to research integration. *JAMA* 1988; 259:1685-9.
17. Egger M, Davey Smith G. Meta-analysis. Potentials and promise. *British Medical Journal* 1997; 315:1371-1374.
18. Freiman J, Chalmers T, Smith H, Kuebler R. The importance of beta, the type II error and sample size in the desining and interpretation of the randomized control trial: survey of 71 "negative trials". *N Engl J Med* 1978; 299: 690-694.

PhUSE 2007

19. Egger M, Davey Smith G. Meta-analysis software. BMJ 1998; 7126 Vol 316 (<http://www.bmj.com>)
20. Parmar KB, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. Stat Med. 1998; 17: 2815-2834
21. Zamora J, Abaira V, Muriel A, Khan K, Coomarasamy A. Meta-Disc: a software for meta-analysis of accuracy data. 2006; 6:31
22. Tinazzi A, Carinci F, Torri V. Graphical aspect of meta-analysis using SAS/GRAPH and its specific annotate language. SEUGI 1996, Hamburg
23. Wang M, Bushman B. Integrating results through meta-analytic review using SAS Software. Cary, NC, USA; SAS Institute 1999
24. Dimakos I. Using SAS Software to perform Meta-Analytic integrations. SUGI 1996.
25. Tierney J, Tinazzi A, Stewart L, Burdett S. A demonstration and guide to SCHARP, a software application for the analysis and plotting of individual patient data (IPD) meta-analysis. The Cochrane Colloquium. Dublin, 2006. Workshop 43

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Angelo Tinazzi

SENDO Tech S.r.l.

Via Visconti di Modrone 12

Milan / 20122

Work Phone: 0039 02 76420426

Fax: 0039 02 76017484

Email: tinazzia@sendo-org.it

Web: <http://www.sendofoundation.org>

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.