A Clinician’s Guide to Understanding, Interpreting and Evaluating a Meta-analysis

* Dr Binny Krishnankutty, Gijo Thomas, Dr Salini Sasidharan

Most of the clinicians find it difficult to interpret a meta-analysis study and are unsuccessful in the critical evaluation of the publication. Why is it so? Firstly, meta-analysis is not taught in the undergraduate curriculum unlike randomized controlled clinical trials, which is introduced in preventive medicine classes. Hence we often lack insight on the meta-analysis process. Secondly, meta-analysis is a complex analysis conducted on results obtained from multiple clinical trials and the results are published with jargon which appears Greek and Latin to many of the clinicians. Thirdly, our exposure to meta-analysis studies is limited compared to clinical trials. This article would enable doctors understand the common terminologies used in meta-analysis and help them evaluate and interpret the results of a meta-analysis study.

What is a Meta-analysis?

There is a common situation faced by most of the clinicians practicing Evidence Based Medicine. A literature review conducted on a particular question leads to many clinical trials which often show contradicting results. There is a dilemma as to which one to believe. Evidence-based medicine has introduced well-defined rules for the critical evaluation of medical data. All studies will be evaluated on its merits and demerits and all the ‘not-so-good’ studies will be excluded. Still we may have studies with opposite results. At this juncture we would like to gauge the direction of the overall effect: how much evidence is there in the data from these contradicting studies to favour one outcome over the other, and we would like to have an estimate of the size of the effect. This is where a meta-analysis comes to help.

Meta-analysis is an analytical method where both independent and different studies are integrated and their results pooled into a single common result. When compared to controlled clinical trials, meta-analysis has the advantage that it gives an unbiased interpretation of results with the advantage of being less influenced by the personal opinion of the researcher. Hence meta-analysis is considered as the highest level of medical evidence in the hierarchy of research evidence (Figure 1).

Figure 1: Hierarchy of evidence

* MBBS, MD (Pharmacology), PGDCR
Manager, Global Medical Affairs,
Dr. Reddy’s Laboratories Ltd., Ameerpet,
Hyderabad – 500016
Mobile: 09949833055
The term meta-analysis was coined by Gene V Glass in 1976 who defined meta-analysis as the ‘statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings’. Meta-analysis is usually done on the results reported from other clinical trials. However, it may also be done on individual patients’ data (called meta-analysis of individual patients). Meta-analysis helps in the following:

- Improves the assessment of the outcome by estimating the effect of treatment, by combining results of contrasting studies
- Improves the statistical power of the comparison and answers new questions
- Helps in analyzing trends and also helps in the analyses of sub-groups (this may not have adequate statistical power in individual studies) by including these from multiple studies

Planning a Meta-Analysis: the background information

There are various steps in planning a meta-analysis. Like any other research study, meta-analysis also starts with a research question for which a definitive answer is being sought.

1. Initially, the outcome of interest to be studied in the meta-analysis should be defined. This may be a clinical outcome like death, occurrence of cancer, complete cure, efficacy of a treatment or a composite outcome (for example, all gastrointestinal complications which include, upper GI bleeding, lower GI bleeding, occult bleeding, symptomatic gastroduodenal ulcers or erosions, obstruction, or perforation). Is it a valuable outcome relevant to the indications and drugs being studied?

2. Next step in this process is procuring relevant literature. One should have a protocol with well-defined criteria for searching and identifying the appropriate publications. The source of literature search should also be mentioned. The commonly used medical bibliographic sources are Pubmed/Medline, Embase, Index Medicus and sometimes, Google Scholar too. Often, conference literature is also used for conducting meta-analysis. But this has got inherent shortcomings with respect to data reliability and publication bias. Have the authors clearly stated the source of bibliography? Is it an authentic source of medical literature? Are the authors missing out valuable literature by choosing one/two bibliographic databases where only limited information will be available? Did they search with multiple search key-words to extract all possible publications?

3. A clearly documented literature evaluation and selection process should be judiciously followed by the authors. Have the authors selected the right patient population? Is the eligibility criteria same in all the selected studies? Have the authors used any scoring system (quality score) for evaluating the merits and demerits of each clinical trial publication before selecting them? Have the authors modified the inclusion criteria of some of the studies to enable the study to fit into the selection criteria? In a meta-analysis article, the authors should clearly mention all these details. What was the design of the selected clinical studies-were they double blind randomized controlled clinical trials, or were they mostly observational studies?
Bias in Meta-analysis

Most often bias could creep into meta-analysis as this is a secondary evaluation of clinical trial results. Different types of biases can alter the results of a meta-analysis. Most importantly, publication bias and selection bias affect the meta-analysis outcomes. Usually, studies with a positive outcome showing a significant difference are published while studies with a negative outcome (where there was no statistically significant difference shown) are not published. When published studies alone are selected for meta-analysis, the authors are inheriting publication bias also into the meta-analysis. There is a possibility that the results of the meta-analysis might be different had they included the results from the unpublished studies. This is a tricky situation. But how can the authors get the results of unpublished studies? Results of some of the unpublished studies may be presented in conferences or in scientific discussions. These should also be included. Another good source is clinical trial registries (www.clinicaltrials.gov, www.ctri.in) where clinical trial details are registered. The contact details of people doing clinical trials in the concerned topic of investigation can be obtained from these sites and they can be directly contacted for obtaining the results. In some cases, results of individual clinical trials can be obtained from the pharmaceutical companies sponsoring those clinical trials.

Whatever precautions are taken by the authors of the meta-analysis, publication bias might affect the study outcome. The QUOROM (Quality of Reporting of Meta-analyses) statement recommends the inclusion of a flow chart showing the information about the number of randomized controlled trials (RCT) identified, included and excluded from meta-analysis (Figure 2). Additionally the authors should evaluate the presence of publication bias, since its effects cannot be completely eliminated. The authors should also include comments in the discussion section of meta-analysis article stating whether the results may have been influenced by publication bias.

**Figure 2: Flow-chart showing study selection process**

There are many methods for evaluating publication bias. The most commonly used is the ‘Funnel Plot’, a graphical method for detecting publication bias, introduced by Light and Pillemer in 1984 (Figure 3). This plot is drawn using the effect size and the sample size. The X axis may represent risk ratio, odds ratio, risk difference or the treatment difference. The Y axis may represent sample size, variance or other options. In the funnel plot shown here, the sample size of selected studies are included in the Y axis and Risk ratio is included in the X axis. Individual studies are represented by ‘+’ marks in the graph. In the absence of publication bias, the diagram looks like an inverted funnel with studies symmetrically placed inside the funnel. Figure 3
shows the funnel plot in the absence of publication bias. One important thing to be noted here is that asymmetrically placed studies inside the funnel need not mean that publication bias is present. However a gross asymmetry (most of the studies falling in one half of the funnel) indicates presence of publication bias. Another important point to be kept in mind while evaluating a funnel plot is that funnel plot is useful only if large number of studies are included in the meta-analysis. It should always be remembered that this is a very approximate method since it gives only a visual judgment of the data. There are other statistical methods also for estimating the asymmetry in funnel plots and to quantify the publication bias in the study (for e.g., Begg’s rank correlation method, Klein’s method).

Figure 3: Funnel Plot

What is Heterogeneity?

In meta-analysis we are combining different studies and pooling the effect. Homogeneity in a meta-analysis means that the results of each individual trial included are mathematically compatible with the results of any of the other studies. Heterogeneity, on the opposite, is the measure of variability across studies. The study protocols of individual studies may differ from each other and this difference may lead to heterogeneity. But even if the same protocol is used in all studies, variability in study quality, possibly due to the mistakes in implementing the protocol or due to variations in the patient population, may give rise to heterogeneity. It is imperative to include tests for heterogeneity in the meta-analysis. These tests will show whether there is heterogeneity in the results. Broadly speaking, heterogeneity can be guessed from the appearance of confidence intervals of individual studies shown in the Forest Plot (vide infra). If the confidence limits of the individual studies do not cross that of each other, there is high chance for heterogeneity in the study.

Tests for heterogeneity are helpful in deciding whether to adopt the fixed effects model or the random effects model in calculating the overall treatment effect. Test of heterogeneity is a variant of the chi square test and the most frequently used statistic is Cochran’s Q. If the p value of the test of heterogeneity is 0.05 or less (heterogeneity present), random effects model will be used and if the p value is more than 0.05 (heterogeneity absent), fixed effects model will be used. A fixed effects model is more robust than the random effects model as there can be some inflation in the confidence interval in the random effects model. Hence it is better to approach the results with caution if there is heterogeneity and random effects model is used in pooling of treatment effects.

Cochran’s Q test has very low statistical power and hence the alpha value (type I error) value should be limited to 0.10 (or 10%). Additionally, Cochran’s Q estimates all heterogeneity (heterogeneity among studies and within individual studies) and not the true heterogeneity alone, i.e., true difference among studies. Hence $F$ statistic is
used additionally. $I^2$ describes the percentage of total variation across studies due to heterogeneity. $I^2$ is a measure to quantify the heterogeneity. It is arbitrarily chosen that an $I^2$ value of 25%, 50% and 75% indicate mild, moderate and severe heterogeneity, respectively.

**Interpretation of Meta-analysis results (Forest Plot)**

Figure 4 represents the Forest plot showing the final results of meta-analysis. This shows the pooled odds ratio or risk ratio calculated from that of each study. There is a standard format followed in Forest plots. On the left hand side, all studies included (sometimes, names of first authors are put) in the meta-analysis are listed in ascending chronological order. Additionally it should give the year of publication, number of events/number of subjects included, odds ratio (or risk ratio) of individual studies with their confidence limits, the p-value of each study and a horizontal line with a blob somewhere in the middle (square blobs). The last row in the plot will show the data for the pooled analysis (diamond shaped blob).

**Figure 4: Forest Plot showing pooled analysis**

<table>
<thead>
<tr>
<th>Study name</th>
<th>Risk ratio Lower limit</th>
<th>Upper limit</th>
<th>Z-Value</th>
<th>p-Value</th>
<th>ARB</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIFE (2002)</td>
<td>1.115</td>
<td>0.964</td>
<td>1.289</td>
<td>1.465</td>
<td>0.143</td>
<td>356/4605</td>
</tr>
<tr>
<td>TROPHY (2006)</td>
<td>1.316</td>
<td>0.207</td>
<td>5.844</td>
<td>0.362</td>
<td>0.718</td>
<td>4/306</td>
</tr>
<tr>
<td>TRANSCEND (2008)</td>
<td>1.184</td>
<td>0.172</td>
<td>1.164</td>
<td>0.099</td>
<td>0.253</td>
<td>236/206</td>
</tr>
<tr>
<td>ON-TARGET (2008)</td>
<td>1.041</td>
<td>0.145</td>
<td>1.147</td>
<td>0.011</td>
<td>0.417</td>
<td>742/554</td>
</tr>
<tr>
<td>PROFESS (2009)</td>
<td>0.962</td>
<td>0.121</td>
<td>1.117</td>
<td>0.510</td>
<td>0.610</td>
<td>326/10016</td>
</tr>
<tr>
<td></td>
<td>1.056</td>
<td>0.088</td>
<td>1.128</td>
<td>1.612</td>
<td>0.107</td>
<td>340/10048</td>
</tr>
</tbody>
</table>

The black line down the middle is the 'line of no effect'. The line passes through a risk ratio of 1.0. One side of the central line shows that ARB is worse while the other side shows that the control drug is worse. If the confidence interval of the result (the horizontal line) crosses the line of no effect (the vertical line), that can mean either that there is no significant difference between the treatments or that the sample size was too small to allow us to be confident where the true result lies. If the confidence interval of the pooled result lies entirely on one side of the central line, it means that there is a significant response and the side on which it lies indicates whether ARB is worse or the control drug is worse. But in our case, the diamond-shaped blob is crossing the central line. Moreover, the p-value of the pooled
results is 0.107 indicating that there is no significant difference in new cancer occurrence between ARBs and other drugs/placebo (confidence interval of risk ratio was 0.988 to 1.128).

Conclusion

This article gives an overview of meta-analysis processes, methods for evaluating the methodology adopted, common terminology mentioned in the meta-analysis articles and the interpretation of results and plots. Since meta-analysis is a secondary analysis, results might completely reverse on including studies by using different selection criteria. Hence caution should be exercised when a study is evaluated for its merits and demerits and the results are interpreted. Critical review of the meta-analysis article should focus on the accuracy of data, methodologies adopted, evaluation of the merits and demerits of the inclusion/exclusion criteria used in the individual studies, selection criteria used by authors for the inclusion of studies in the meta-analysis, presence of subgroup analyses and the genuineness of interpretation of results. Moreover, the reviewer should also check whether heterogeneity and publication bias aspects have been taken care of or not. Being the highest level of medical evidence, meta-analysis remains an invaluable tool in answering the research question by pooling results from contradicting RCTs, if done properly. Similarly it has the power to mislead the scientific fraternity and create havoc, if wrongly used.

References

1. Merlin T, Weston A, Tooher R. Extending an evidence hierarchy to include topics other than treatment: revising the Australian ‘levels of evidence’. BMC Medical Research Methodology, 2009; 9: 34.
