Journal of Parenteral and Enteral Nutrition
Volume XX Number X
Month 201X 1–8

DOI: 10.1177/0148607116661841
jpen.sagepub.com
hosted at
online.sagepub.com

Understanding Systematic Reviews and Meta-Analyses

Ronald L. Koretz, MD[1]; and Timothy O. Lipman, MD[2]

Abstract

Systematic reviews should be distinguished from narrative reviews. In the latter, an editor asks an expert to sum up all of the information that is known about a particular topic. However, the expert is under no constraints regarding what he or she does, or does not, choose to include in the review. As a result, his or her bias can influence the final message. A systematic review, which may or may not be written by experts, typically asks a narrower question, and then answers it using the entirety of the medical literature. The systematic review process includes computer searches to identify the pertinent literature, a statement of the inclusion and exclusion criteria for identified studies, a list of items of interest to extract from each study, a method to assess the quality of each study, a summary of the evidence that has been found (which may or may not involve attempts to combine data), a discussion of the evidence and the limitations of the conclusions, and suggestions for future research efforts. If the data are combined, that process is called meta-analysis. In meta-analysis, an estimate of the reliability of each study is made, and those that appear to be more reliable are weighed more heavily when the data are combined. While systematic reviews depend on a more preplanned method and thus, unlike narrative reviews, contain sections on method, they can be easily read once the reader becomes familiar with the vocabulary. *(JPEN J Parenter Enteral Nutr. XXXX;xx:xx-xx)*

Keywords

systematic review; meta-analysis; evidence-based medicine; critical reading

In the era of evidence-based medicine, a systematic review and meta-analysis is generally considered the gold standard. In this article for the *JPEN* reader, we will attempt to explain how to read and understand systematic reviews and meta-analyses. We plan to approach this tutorial by answering the following questions:

- What is a systematic review?
- What is a meta-analysis?
- Is a meta-analysis the same as a systematic review?
- What are the components of a systematic review?
- What are the components of a meta-analysis?
- How can a systematic review be judged?

What Is a Systematic Review?

To answer this question, one needs to understand the difference between a narrative review and a systematic review. Any review, obviously, attempts to summarize the body of knowledge about a particular topic.

In a narrative review, an "expert" summarizes the important aspects concerning a particular topic. It is assumed that this expert will be objective in presenting the pertinent information. Unfortunately, bias, usually unintentional, is often a problem. Mulrow[1] proposed a set of processes that could be employed in narrative reviews to maintain objectivity. These include the following:

- A clear statement of the purpose of the review
- A description of the methods used to identify the type of evidence that was being sought
- Inclusion and exclusion criteria to define whether any piece of evidence would be used
- A description of the method that was used to assess the quality of each piece of evidence

From the [1]Olive View–UCLA Medical Center, Sylmar, California, USA; and the [2]Gastroenterology, Hepatology, and Nutrition Section, Veterans Affairs Medical Center, Washington, DC, USA.

Financial disclosure: RLK acknowledges the ongoing support of GIIssues, Inc, a 501 (c) (3) nonprofit organization dedicated to the practice and dissemination of evidence-based medicine. GIIssues, Inc supports educational expenses for the author as well as any expenses related to the creation of papers and other educational products. GIIssues, Inc does not provide any salary support to the author and did not provide any support for this particular article. There are no research materials that are related to this article that can be accessed other than the stated references.

Conflicts of interest: None declared.

Received for publication March 29, 2016; accepted for publication June 5, 2016.

Corresponding Author:
Ronald L. Koretz, MD, Department of Medicine, Olive View–UCLA Medical Center, 14445 Olive View Dr, Sylmar, CA 91342, USA.
Email: rkoretz@msn.com

- A synthesis of the evidence (either qualitative or quantitative)
- A discussion of the limitations of the evidence
- A summary of the evidence
- Suggestions for areas where future research is needed

When this group assessed a series of narrative reviews in the mid-1980s, they reported that most of these papers lacked evidence for the use of these processes.[1] A decade later, the same problem was still present.[2] The reader can never be certain in a narrative review that all relevant literature has been considered and that he or she is reading an unbiased summary of available knowledge. Most frequently, a narrative review reflects the reviewer's opinions about the topic.

Systematic reviews differ from narrative reviews. The systematic review assimilates information about a particular topic or question with more rigor, sophistication, and (most importantly) transparency. It incorporates all of the processes noted in the above bullet list. While content expertise can be helpful, systematic reviews need not be performed by experts. A systematic review is a formal process to gather and evaluate literature to answer a defined question. The systematic review process is transparent. All of the steps, beginning with the posing of the question, continuing with the plans for describing the inclusion criteria of trials, searching for pertinent articles, extracting the necessary data from each one, assessing those articles for methodologic quality, contacting the original investigators for missing information, deciding if and how to combine the data, and concluding with the analyses of the information obtained, are laid out in a methods section. This is done ideally with a protocol that was written before the actual review was begun. Systematic reviewing is, in many aspects, very similar to doing a research project. Thus, any reader could duplicate the method, find the same literature, and derive similar conclusions.

We will largely focus on systematic reviews of randomized clinical trials in the material that follows. However, it should be remembered that one can use the same principles to undertake systematic reviews of observational studies.

A recent systematic review that assessed the method of published systematic reviews of Mediterranean diets suggested that, at a minimum, a systematic review should be defined as a systematic search, which extracts information from studies following an a priori protocol.[3]

What Is a Meta-Analysis?

There are 2 types of systematic reviews, qualitative and quantitative. Qualitative reviews discuss the studies that are found but make no effort to combine the data from them; quantitative reviews combine the data. There are 2 major reasons to combine data from different studies. The first is to increase the ability to see a difference between 2 groups, thereby reducing the chance of having a type II error (ie, missing the existence of a true difference). The second is to increase the precision of the estimated effect. Data combination can be performed by simply adding the numerators and denominators from each study together. Meta-analysis is a more sophisticated method for combining data; it assigns a weight to each study depending on how reliable each study appears to be from a statistical perspective and then adds the weighted numbers together. Thus, while the term *meta-analysis* is intimidating to many people, it fundamentally represents a sophisticated process of addition.

Two different kinds of data can be pooled, dichotomous and continuous. Dichotomous data reflect outcomes that either did, or did not, occur (eg, death, infection). Continuous data reflect outcomes that are specific numbers in a continuum (eg, blood pressure, days in the hospital). The effect of an intervention on a dichotomous outcome can be expressed in 3 ways. The risk ratio is the percentage of patients in the intervention group who had that outcome divided by the percentage of patients in the control group who had that outcome. The odds ratio is the odds of the outcome happening in the intervention group divided by the odds of it happening in the control group. Finally, the risk difference is the difference in the percentages between the outcome in the intervention group and that in the control group. A continuous outcome is presented as a mean difference, which is the differences between the average (or mean) of the outcome in the intervention group and the mean of the outcome in the control group.

The previously noted review of systematic reviews defined a meta-analysis as, at a minimum, a provision of a weighted effect size for an outcome of interest.[3]

Is a Meta-Analysis the Same as a Systematic Review?

A systematic review is not the same thing as a meta-analysis. The former represents an effort to obtain all of the available information to answer a clinical question. The latter is a mathematical manipulation. One can perform a systematic review without doing a meta-analysis, and one can meta-analyze data without doing a systematic review.

What Are the Components of a Systematic Review?

The systematic review process begins with identifying a specific question and writing a protocol. As we just noted, this protocol describes explicitly the strategy for the literature search, the inclusion and exclusion criteria for trials to be used, the time period of the search, the data abstraction sheet, the specific data analyses to be done (including potential plans to combine the data or not, criteria for quality assessment of each included study, and a list of subgroup or sensitivity analyses [defined and discussed in the section on the components of meta-analysis below] to be performed).

What should the reader look for and expect from a systematic review? A number of features are usually present.

There should be evidence of a written protocol. Ideally, such a protocol should be published separately. Even if a published protocol is not available, there should be evidence in the methods section of the systematic review that thought was put into the following factors:

- What is the question being addressed?
- What type of study was going to be assessed (eg, randomized clinical trial, observational study)? Were there language restrictions (eg, English only, English and other languages only, any language)? Were only full published studies acceptable, or would the review also include abstracts, doctoral theses, studies that were listed on trial registries, and other reports of research endeavors that are not typically published as full articles in medical journals (and referred to as "gray literature")? What electronic databases or other sources for references were searched and what search terms were employed? What dates were searched? What inclusion and exclusion criteria were employed for each identified trial? What data were to be sought in each reference? What was the primary outcome measure to be assessed, and what secondary outcomes and subgroup/sensitivity analyses (see the section regarding the components of a meta-analysis below) were planned?
- What effort was made to assess the quality (otherwise known as the "risk of bias") of each reference used in the review? We know that there are several ways in which bias can creep into randomized trials. Probably more often than not this occurs at a subconscious level by the investigator and is not a purposeful act. It is easy to understand how the lack of blinding can lead to misinterpretations or overestimations of benefit.[4] However, there are other domains of risk of bias, including generation of the allocation sequence,[5] concealment of allocation[6] (knowing ahead of time into which arm of a randomized trial a particular participant will be assigned), attrition bias[7] (loss of patients during the trial), selective outcome reporting[8] (often occurring when differences that are not significant are not reported in the study), early stopping of trials for perceived benefit,[9] and vested (either economic or academic) interests. Trials containing these various risks of bias will tend to overestimate benefit and underestimate harm.[6,10]
- Is there a plan to assess for publication bias? Publication bias is a phenomenon that occurs when particular studies are performed but not published; this will be discussed in more detail below.
- Was there a plan to pool the data? If so, how and what planned analyses were going to be done?

Usually, there should be a figure that shows the total number of studies that were initially identified and how those identifications occurred, the number of duplicate references, the number of studies that were considered for a more in-depth review after the title and abstract identified from the computer search were assessed, and the final number of studies that were included. Such figures are known as "PRISMA flow diagrams,"[11] and a fictitious example of one is demonstrated in Figure 1. (PRISMA is an acronym for Preferred Reporting Items for Systematic reviews and Meta-Analyses.)

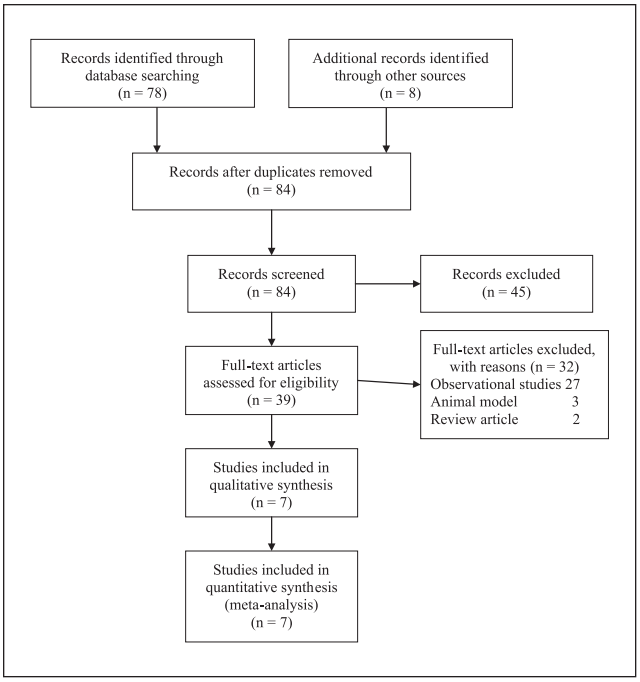

Figure 1. PRISMA flow diagram for systematic review on the use of dental abrasives in treating odontopruritus. PRISMA flow diagram from Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(6):e1000097.

The method for abstraction of data should be described. At a minimum, 2 authors should abstract data, and the method for resolution of conflicts needs to be described. Mistakes are easy to make in data abstraction, but it is very unlikely that the same mistake would be made by 2 people.

The identification of all pertinent trials is not a straightforward exercise. If only MEDLINE is searched, at least half of the pertinent literature may be missed.[12–16] To find all of the trials, multiple databases should be searched, relevant investigators should be asked about their knowledge of other trials, reference lists of identified papers should be searched, and trial registries should be assessed. Ideally, hand-searches of pertinent journals should be conducted to identify abstracts from meetings. Searching registries and looking for abstracts are techniques that are employed to identify the gray literature. Whenever pertinent information is not available in the published manuscript or trial registry, the systematic reviewers should try to contact the original investigators.

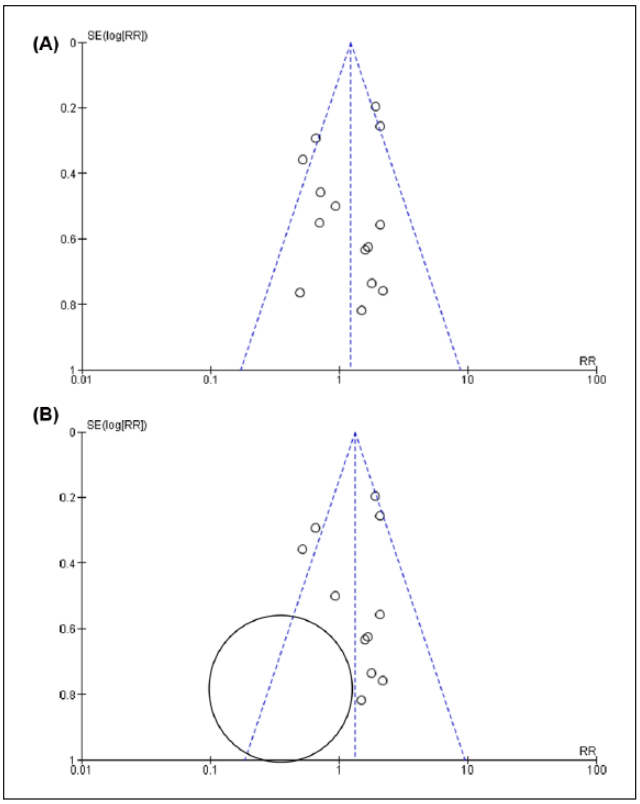

Figure 2. (A) Funnel plot analysis failing to show publication bias. (B) Funnel plot analysis suggesting publication bias. RR, risk ratio; SE, standard error.

If a study is not identified by the individuals doing the systematic review, it will not be included in that systematic review. Unfortunately, studies are sometimes completed but never reported as full articles or even as abstracts. Such studies may go unreported for a variety of reasons, including null findings, small sample size, findings contrary to expectations, study performed by a student who has subsequently moved and/or never got around to writing and submitting the article, or a rejection of the manuscript by medical journals. Since trials (especially smaller ones) that fail to show differences are less likely to be published, the failure to identify them can result in an overestimation of the true effect.[17] This phenomenon is termed *publication bias*.

Obviously, the authors of a systematic review cannot assess what has not been published. However, techniques can be used in an effort to detect publication bias. One commonly employed technique to identify publication bias is to do a "funnel plot analysis."[18] This is a graphic representation of all of the trials; each trial is represented by a single point that is located on the graph such that the observed effect is on the x-axis and some measurement of the size or precision of the trial (typically the standard error) is on the y-axis. Larger and/or more precise trials are likely to produce estimates of effects that are closer to the truth. A fictitious example of funnel plot analysis is depicted in Figure 2.

In clinical trials, smaller/less precise trials will be located lower down on the y-axis, and there will be a wider spread in the estimated effect. However, if publication bias is present, the funnel plot would be expected to show a hole on one side or the other (where the missing trials should be). In Figure 2A, all of the (fictitious) trials for the treatment of odontopruritus (itchy teeth) are represented, including a number of small trials in which the effect varied over a wide range. In Figure 2B, 3 of those small trials, all failing to find a significant effect, were removed from the analysis. As can be seen in Figure 2B, there is a region in the lower left area of the funnel plot (demarcated by the circle) where those trials are missing.

Funnel plot analysis needs a large number of points to be able to detect the holes; it is generally recommended that it only be undertaken if there are at least 10 trials available.[19] A classic example of publication bias that occurred in real life was an assessment of the risk of cancer in Barrett's esophagus.[20]

Although not reflecting the quality of a systematic review or meta-analysis, there has been a recent suggestion that systematic reviews and meta-analyses may be "ghost-written" to enhance the curriculum vitas of academic scientists.[21] While the average reader cannot be expected to recognize such ethically dubious publications, the reader should be aware of the existence of this possibility.

Table 1 provides a summary of considerations that Huedo-Medina et al[3] employed when they assessed the methodologic quality of a number of systematic reviews that addressed Mediterranean diets.

What Are the Components of a Meta-Analysis?

As we have already noted, data combination (either data pooling or meta-analysis) provides 2 important advantages: increased ability to see smaller differences (power) and increased precision. However, we have to be concerned about combining apples and oranges—how different are various trials, and are they similar enough to even consider combining? We call these differences in trials *heterogeneity*.

No 2 trials are ever exactly the same, so there will always be some heterogeneity. When one considers a plan to combine data from different trials, the first question that should be asked is whether it makes any clinical sense to do so. If it is clinically nonsensical, then data combination should not be done. However, if it does make clinical sense to combine data, various statistical tests and maneuvers are available to assess for statistical heterogeneity. These tests have limitations. They are not very sensitive, and they depend only on the actual numbers, not taking the study design into consideration. As an example, imagine a research center in which 4 randomized clinical trials were undertaken to assess the effect of some intervention on mortality, but the actual trials looked at totally different interventions in totally different disease states. If the mortality rates

Table 1. Quality Considerations for Systematic Reviews and Meta-Analyses.

Category	Component
Protocol	Prepared before review undertaken
	Clear description of data to be abstracted
Literature search	Results of search assessed by at least 2 people
	At least 2 people independently abstracted data from each study
	Search included multiple databases
	Search technique described in sufficient detail that it could be reproduced
	Efforts made to identify unpublished studies
	List of included studies provided
	List of excluded studies, as well as reason(s) for exclusion, provided
Description of studies	Study characteristics presented
	Methodologic quality (risk of bias) of each included study assessed and reported
Data analysis	If data were combined, the method of pooling that was employed was correct
	Subgroup/sensitivity analyses performed to address sources of heterogeneity, including risks of bias
	Statistical justification for the effect measurement (eg, odds ratio) employed
	Publication bias was addressed formally
Conflict of interest (COI)	All potential COIs provided

Concepts from Huedo-Medina TB, Garcia M, Bihuniak JD, et al. Methodologic quality of meta-analyses and systematic reviews on the Mediterranean diet and cardiovascular disease outcomes: a review. *Am J Clin Nutr*. 2016;103:841-850.

in the treatment and control groups happened to be similar, a statistical test for heterogeneity would indicate that heterogeneity was not present.

Depending on the presence or absence of heterogeneity, different statistical models are employed to combine data with meta-analysis. If it is believed that the studies are homogeneous, the "fixed-effect" model is employed. This mathematical model assumes that all of the trials evaluated patients who came from the same general population. The "random-effects" model, on the other hand, is used to deal with heterogeneity and assumes that the study populations themselves are different.

The presence of heterogeneity actually can be useful, as it provides an opportunity to focus on how or where the intervention might be most effective. Systematic reviewers should consider in advance (when they write their protocol) what heterogeneous factors might be present and important. In this way, they can plan to perform subgroup (which become hypothesis-generating exercises) or sensitivity (which assess the robustness of the conclusions) analyses that will account for these factors. Subgroup analyses divide the trials by a priori identifiable characteristics (eg, the duration of treatment) and then assess if the presence or absence of that characteristic enhanced or lessened an effect (thus creating a testable hypothesis). One commonly used sensitivity analysis is to divide the trials into those at high risk of bias and those at low risk of bias and then separately analyze each of these 2 smaller groupings of trials. If an effect is only seen in the meta-analysis of the trials at high risk of bias, there is cause to worry that an observed benefit is actually attributable to methodological and/or reporting deficiencies and not to a true effect of the intervention.[22]

Meta-analyses are usually graphically presented in figures called forest plots. An example of one such plot is presented in Figure 3A. In this meta-analysis, 7 (fictitious) trials assessed the value of various types of dental abrasives in treating odontopruritus. The first 7 rows represent the individual trials; these have been listed by the type of abrasive, although more typically, each trial would be listed by the first author and date of publication of the study. The next 4 columns represent the numbers of patients who were relieved by the intervention, the total number of patients in the intervention group, the number of patients who were relieved in the control group, and the total number in the control group. The sixth column is the calculated weight of each trial. The seventh column represents the individual risk ratio (RR) for each trial.

The figure to the right in our Figure 3A example illuminates the individual data for each trial: the location of the box is the RR for that trial, the size of the box reflects the weight of that trial, and the horizontal line presents the 95% confidence interval (CI) for that trial.

The eighth row shows the total numbers of patients and the calculated estimated effect (ie, the RR) for the meta-analysis. In our example (Figure 1A), the RR is 1.68 (the RR being larger than 1.0 means that more patients in the treated arm achieved relief), and the 95% CI for the meta-analysis does not cross the line of equivalence (the vertical line at 1.0 in the diagram). (The failure of the 95% CI to cross the line of equivalence defines "significance" in the same way that a P value <.05 leads investigators to declare a difference as being significant.) In the figure, the diamond summarizes this calculation: the vertical line between the high and low points of the diamond is the estimated RR, and the right and left sides of the diamond are the 95% CI. In other words, the estimated effect from this "fixed-effects" model meta-analysis suggests that individuals using dental

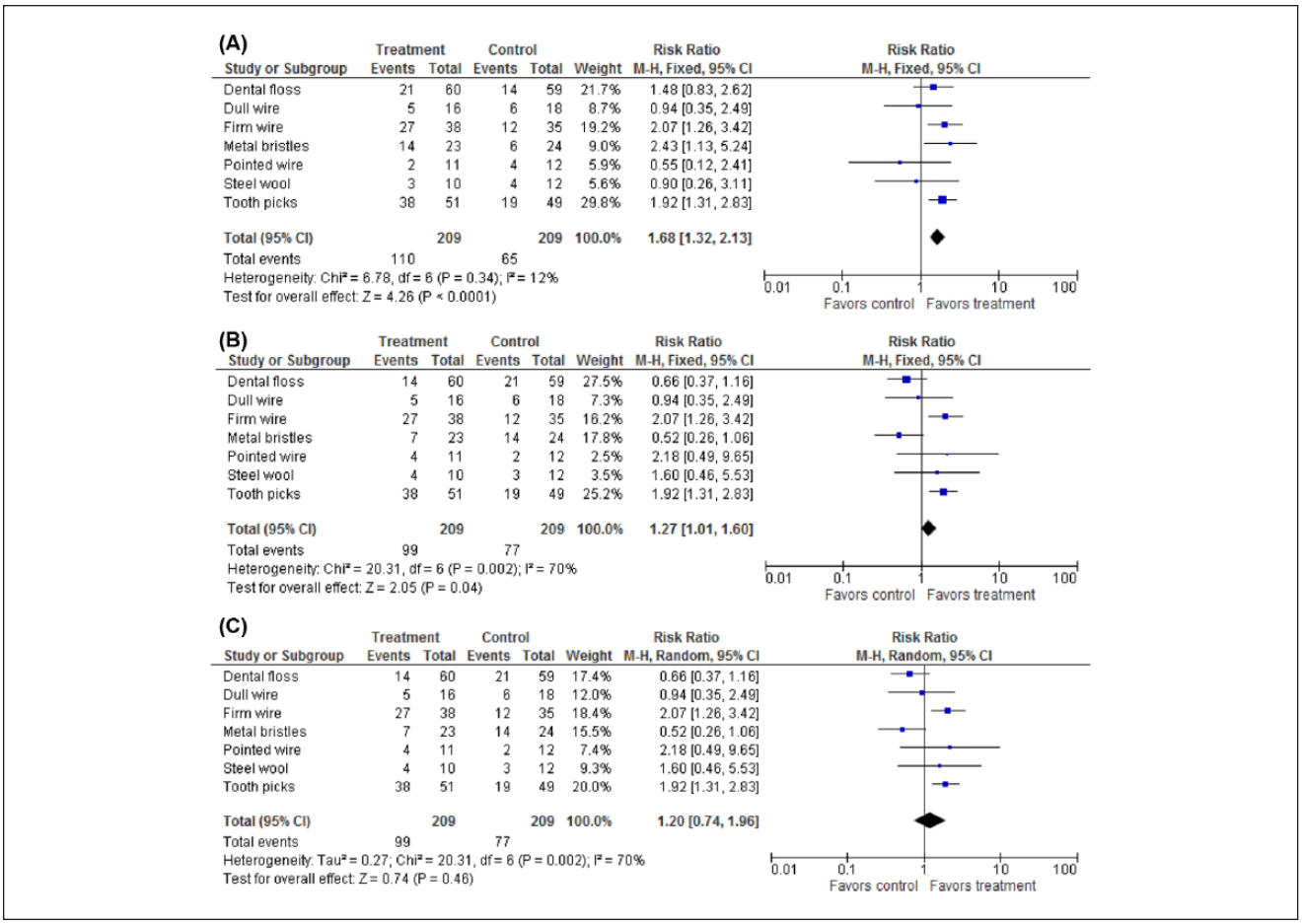

Figure 3. (A) Forest plot with homogeneous trials (fixed-effect model). (B) Forest plot with heterogeneous trials (fixed-effect model). (C) Forest plot with heterogeneous trials (random-effects model). CI, confidence interval; M-H, Mantel-Haenszel.

abrasives had complete relief 1.68 times more often than did those in the control group.

The ninth row is the total number of individuals achieving complete relief in each group. If one wished simply to perform data pooling, this could be done using rows 8 and 9 (ie, 110/209 [53%] vs 65/209 [31%]).

The 10th row presents the statistical tests for heterogeneity. Usually, heterogeneity is suspected when the *P* value is <.10; in our example, $P = .34$. (The actual statistical test is a χ^2 one, and the *P* value depends on that value as well as the number of degrees of freedom [or *df*].) The I^2 statistic is an estimate of how much of the differences in the various trials is more than what could be expected by chance; in this example, the number is small (12%). I^2 values >50% suggest the presence of important heterogeneity.[23] The *z* score for overall effect (11th row) is a more complicated statistical calculation that addresses the probability that the difference between the 2 groups is due to chance.

This meta-analysis (our fictitious example in Figure 3A) was purposely created with numbers that would demonstrate homogeneity. The model that was used (and can be seen on the forest plot by looking at the top of the diagram and seeing the word *Fixed*) was the fixed-effect model. We can create a situation with manipulation of the data such that the trials are heterogeneous (Figure 3B). When the fixed-effect model is used with these new data, a "significant" effect is still present, albeit smaller: those using the abrasives had complete relief 1.27 times more often, and the 95% CI (1.01–1.60) does not overlap the line of equivalence. However, as can be seen in row 10 (presenting the tests for heterogeneity), the *P* value is now .002, and the I^2 is 70%; both numbers indicate substantial heterogeneity.

When one sees these types of results, it would be more appropriate to use the random-effects model. This model gives more weight to the smaller trials and, as a result, typically increases the size of the 95% CI. The meta-analysis employing the random-effects model is depicted in Figure 3C. In this illustration, the data are the same as in Figure 3B. While the estimated effect has not changed very much (RR, 1.20), the 95% CI has widened (0.74–1.96); statistical significance is no longer present.

How Can a Systematic Review Be Judged?

Conclusions drawn from a systematic review cannot be any stronger than the strength of the medical literature from which those conclusions are made. It is for this reason that systematic reviews address the risks of bias in each of the individual references. However, other considerations can also limit the strength of conclusions that are drawn in systematic reviews. A popular approach to the assessment of the strength of evidence in systematic reviews is the GRADE (Grades of Recommendation, Assessment, Development, and Evaluation) method.[24] The GRADE system evaluates a number of factors. These include, in addition to the quality of the evidence, the consistency of the evidence (how similar the estimates of the effect are among all of the individual trials), the directness of the evidence (how important the reported outcomes are to those that would be of interest to clinicians and patients), the precision of the evidence (the width of the confidence interval of the putative effect), and the likelihood of reporting bias. In the GRADE system, meta-analyses of randomized trials begin as a "high" level of evidence ("future research is unlikely to have a substantial impact on our confidence regarding the estimate of the effect"), but the overall assessment can be reduced to "moderate" ("future research is likely to have an important impact on our confidence"), "low" ("future research is very likely to have an important impact on our confidence"), or very low ("the current estimate is very uncertain"). Serious problems with the quality, consistency, directness, precision, or reporting will result in a downgrading from 1 to 3 levels (ie, the high will become moderate, low, or very low).

In conclusion, systematic reviews and meta-analyses are appearing more frequently in medical journals. Systematic reviews and meta-analyses are not the same process but are often confused and the terms used interchangeably. Systematic reviewing refers to a methodologically driven gathering and evaluation of scientific publications, whereas meta-analysis is a form of data pooling or combination. Once one becomes familiar with their structure and content, systematic reviews and meta-analyses should not be difficult to read and understand. It was not our intent to enable the average reader to perform a systematic review or produce the appropriate meta-analysis after reading this article; rather, we have attempted to provide a framework for understanding these important approaches to assessing evidence in health care.

Glossary

Fixed-effect model: The statistical formulation that is employed to perform meta-analysis when it is believed that the individual studies contained participants who all possessed characteristics of a single population

Forest plot: The figure that illustrates the various studies and the final estimated effect of a meta-analysis

Funnel plot: A pictorial demonstration that is commonly used to identify the presence or absence of publication bias

Heterogeneity: The term that is used to describe how different a particular study is from another one or how different is the totality of the studies in a meta-analysis

Meta-analysis: The process of combining data from different studies using evaluations of statistical variability to determine the weight that each study will be assigned in the combination

Narrative review: An article or book chapter that addresses a particular topic and that is authored by one or more experts who use(s) his/her/their own judgment to determine what medical literature needs to be cited

Publication bias: The failure to publish completed clinical studies

Qualitative data synthesis: The presentation of individual studies without any effort made to combine the results

Quantitative data synthesis: The statistical combination of data, usually employing meta-analysis

Random-effects model: The statistical formulation that is employed to perform meta-analysis when it is believed that the individual studies contained participants who did not all possess characteristics of the same population

Systematic review: A process whereby an answer to a defined clinical question is sought by using all of the information available in the medical literature following an established protocol

Suggested Additional Reading Resources

Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315:629-634.

Elashoff JD. Combining results of clinical trials. *Gastroenterology*. 1978;75:1170-1174.

Friedman H, Goldberg JD. Meta-analysis: an introduction and point of view. *Hepatology*. 1996;23:917-928.

Greenhalgh T. How to read a paper: papers that summarise other papers (systematic reviews and meta-analyses). *BMJ*. 1997;315:672-675.

Higgins JPT, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. www.cochrane-handbook.org

Jakobsen JC, Wetterslev J, Winkel P, Lange T, Gluud C. Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. *BMC Med Res Methodol*. 2014;14:120.

Koretz RL. *JPEN* Journal Club 5. Review articles. *JPEN J Parenter Enteral Nutr*. 2014;38:640-642.

Koretz RL. *JPEN* Journal Club 6. Meta-analysis. *JPEN J Parenter Enteral Nutr*. 2014;38:761-763.

Lau J, Ioannidis JPA, Schmid CH. Quantitative synthesis in systematic reviews. *Ann Intern Med*. 1997;127:820-826.

Naylor CD. Meta-analysis and the meta-epidemiology of clinical research. *BMJ*. 1997;315:617-619.

Shojania KG, Sampson M, Ansari MR, Ji J, Doucette S, Moher D. How quickly do systematic reviews go out of date? A survival analysis. *Ann Intern Med*. 2007;147:224-233.

Statement of Authorship

R. L. Koretz and T. O. Lipman equally contributed to the concept/design and acquisition/analysis/interpretation of the material in this manuscript, drafted and critically reviewed the manuscript,

gave final approval, and agree to be accountable for all aspects of the work ensuring integrity and accuracy.

References

1. Mulrow CD. The medical review article: state of the science. *Ann Intern Med*. 1987;106:485-488.
2. McAlister FA, Clark HD, van Walraven C, et al. The medical review article revisited: has the science improved? *Ann Intern Med*. 1999;131:947-951.
3. Huedo-Medina TB, Garcia M, Bihuniak JD, et al. Methodologic quality of meta-analyses and systematic reviews on the Mediterranean diet and cardiovascular disease outcomes: a review. *Am J Clin Nutr*. 2016;103:841-850.
4. Schulz KF, Grimes DA. Blinding in randomised trials: hiding who got what. *Lancet*. 2002;359:696-700.
5. Ioannidis JPA, Haidich A-B, Pappa M, et al. Comparison of evidence of treatment effects in randomized and nonrandomized studies. *JAMA*. 2001;286:821-830.
6. Moher D, Pham B, Jones A, et al. Does quality of randomised trials affect estimates of intervention efficacy reported in meta-analysis? *Lancet*. 1998;352:609-613.
7. Hollis S, Campbell F. What is meant by intention to treat analysis: survey of published randomised controlled trials. *BMJ*. 1999;319:670-674.
8. Chan A-W, Hrobjartsson A, Maahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA*. 2004;291:2457-2465.
9. Montori VM, Devereaux PJ, Adhikari NKJ, et al. Randomized trials stopped early for benefit: a systematic review. *JAMA*. 2005;294:2203-2209.
10. Bassler D, Briel M, Montori VM, et al. Stopping randomized trials early for benefit and estimation of treatment effects: systematic review and meta-regression analysis. *JAMA*. 2010;303:1180-1187.
11. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. *PLoS Med*. 2009;6(6):e1000097.
12. Dickersin K, Hewitt P, Mutch L, Chalmers I, Chalmers TC. Perusing the literature: comparison of MEDLINE searching with a perinatal trials database. *Cont Clin Trials*. 1985;6:306-317.
13. Dickersin K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. *BMJ*. 1994;309:1286-1291.
14. Tran ZV, Weltman A. Differential effects of exercise on serum lipid and lipoprotein levels seen with changes in body weight: a meta-analysis. *JAMA*. 1985;254:919-924.
15. McManus RJ, Wilson S, Delaney BC, et al. Review of the usefulness of contacting other experts when conducting a literature search for systematic reviews. *BMJ*. 1998;317:152-1563.
16. Avenell A, Handoll HHG, Grand AM. Lessons for search strategies from a systematic review, in the Cochrane Library, of nutritional supplementation trials in patients after hip fracture. *Am J Clin Nutr*. 2001;73:505-510.
17. McAuley L, Pham B, Tugwell P, Moher D. Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analysis? *Lancet*. 2000;356:1228-1231.
18. Moreno S, Sutton AJ, Turner EH, et al. Novel methods to deal with publication biases: secondary analysis of antidepressant trials in the FDA trial registry database and related journal publications. *BMJ*. 2009;339:b2981.
19. Sterne FAC, Sutton AJ, Ioannidis HPA, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomized controlled trials. *BMJ*. 2011;342:d4002.
20. Sheehan NJ, Crosby MA, Bozymski EM, Sandler RS. Is there publication bias in the reporting of cancer risk in Barrett's esophagus. *Gastroenterology*. 2000;119:333-338.
21. Filion G. A flurry of copycats on PubMed. http://blog.thegrandlocus.com/2014/10/a-flurry-of-copycats-on-pubmed. Accessed May 12, 2016.
22. Koretz RL, Lipman TO. The presence and effect of bias in trials of early enteral nutrition in critical care. *Clin Nutr*. 2014;33:240-245.
23. Patsopoulos NA, Evangelou E, Ioannidis JPA. Sensitivity of between-study heterogeneity in meta-analysis: proposed metrics and empirical evaluation. *Int J Epidemiol*. 2008;37:1148-1157.
24. Atkins D, Best D, Briss PA, et al, for the GRADE Working Group. Grading quality of evidence and strength of recommendations. *BMJ*. 2004;328:1490-1497.