Meta-analysis in Surgery

Methods and Limitations

Tina T. Ng, MD; Marcia L. McGory, MD; Clifford Y. Ko, MD, MSHS; Melinda A. Maggard, MD, MSHS

The growth of new knowledge continues to advance the surgical disciplines, and several types of literature reviews attempt to consolidate this expansion of information. Meta-analysis is one such method that integrates findings on the same subject from different studies. Within surgery, there is a wealth of literature on a given topic, which needs to be considered collectively. As such, meta-analyses have been performed to address issues like the use of bowel preparation for colorectal surgery and comparisons of outcomes for laparoscopic vs open surgical approaches. A basic understanding of the groundwork required for meta-analysis is fundamental toward interpreting and critiquing its results. This review provides an overview of the principles, application, and limitations of meta-analysis in the context of surgery.

Arch Surg. 2006;141:1125-1130

Meta-analysis refers to the analysis of analyses. I use it to refer to the statistical analysis of a large collection of results from individual studies for the purpose of integrating the findings. It connotes a rigorous alternative to the causal, narrative discussion of research studies which typify an attempt to make sense of the rapidly expanding research literature.

Gene Glass (1976)

Meta-analysis synthesizes the medical literature on a common topic using a rigorous, statistical approach. This method was originally coined by Gene Glass; it was used to help make sense of the growing amount of data in the literature and has become an integral part of the research armamentarium.[1] One of the first medical publications of pooling data was in 1904 by Pearson,[2] and one of the first published, formal meta-analyses was in 1954 by Lasagna et al.[3] Archie Cochrane (1909-1988), a British epidemiology researcher, was one of the foremost leaders in this field and was instrumental in establishing these techniques.[4]

Meta-analysis uses quantitative statistics to combine results from individual studies to determine whether a difference in outcomes exists between 2 treatments or study arm groups. Studies are weighted with regard to their sample size, and overall direction and size of the effect are calculated. Meta-analysis has the potential to resolve controversies within the literature and ultimately shape surgical practice.

Author Affiliations: Department of Surgery, David Geffen School of Medicine at University of California, Los Angeles (Drs Ng, McGory, Ko, and Maggard); Department of Surgery, Greater West Los Angeles Veterans Affairs Medical Center, Los Angeles (Dr Ko).

See Invited Critique at end of article

A recent article on systematic review and meta-analysis by Sauerland and Seiler[5] highlights the important role that these techniques play in evidence-based research for surgeons. They provide groundwork about the steps involved in designing and conducting this type of research. Our review illustrates further the key facets of the method to prepare researchers who are new to performing a meta-analysis but also assists those who want to critique these types of published studies. We describe the detailed methodological steps for meta-analysis, including defining a research question, identifying the selection criteria of studies to include, calculating the odds ratio (OR) and relative

Downloaded from www.archsurg.com at National Institute of Hlth, on October 9, 2007
©2006 American Medical Association. All rights reserved.

risk (RR), and reviewing components of funnel and forest plots. Our overview begins with comparing meta-analysis with other types of literature reviews, followed by the rationale behind performing meta-analysis. We also detail the methods and limitations of this approach, all within the context of surgery.

NARRATIVE REVIEW AND SYSTEMATIC REVIEW: COMPARISON WITH META-ANALYSIS

The narrative review and systematic review both summarize results of previously published literature; however, they have limitations that the meta-analysis attempts to avoid. A traditional "review of the literature" condenses past studies using narrative form and provides generalized conclusions; however, it lacks structure in the inclusion of studies and statistical summary of the data. First, the studies selected may not represent all the available evidence because rigorous selection criteria are not typically used. As such, the summation may support the reviewer's bias. Second, the narrative review does not quantitatively combine the data or weight data, which can create misleading interpretation.[6] Finally, results from different sources can be conflicting and thus a summative review may not reach a statistically justified conclusion.[7,8]

A systematic review is a higher level of review than a narrative approach in that it follows specific methods to define a research question, criteria for study inclusion, and method of data collection.[6,7] Systematic reviews overcome some of the limitations of a traditional review since the prespecified research question and inclusion criteria tend to decrease the amount of bias in the selection of the studies and in the conclusions drawn.

Systematic reviews may or may not formally combine (ie, pool) data or weight the relative contribution of study based on sample size using statistical methods. This method of combining data is called meta-analysis. The method includes attempts to identify studies that are clinically similar enough to allow for qualitative pooling of the data. Tests to identify heterogeneity (ie, wide variations in the effect measure or outcome) are also performed. These terms, *systematic review* and *meta-analysis*, are often confused in the literature. Specifically, a meta-analysis is performed within the context of a systematic review.

Like a systematic review, a meta-analysis follows strict methods to define the study question and establish study inclusion criteria. The major difference is that the meta-analysis approach attempts to identify studies that are similar enough to allow for quantitative pooling of the data. However, if the identified studies are too heterogeneous to allow pooling of the data, then a systematic review may represent the best available method to synthesize the current literature.

META-ANALYSIS: OBJECTIVES

The overarching goal of a meta-analysis is to create an objective conclusion based on the strengths and limitations of the available studies.[6-8] While randomized controlled trials (RCTs) are the gold standard of evidence-based medicine, it is not uncommon for the results from the individual trials to vary, particularly within surgery because of the problems of small sample size and significant variation with respect to procedure techniques, surgeon experience, and postoperative care.

Meta-analysis can help synthesize the results for a number of scenarios where findings of the individual studies show (1) no effect because of small sample size, (2) varying direction of effect, or (3) effect vs no significant effect, all of which are commonly encountered among surgical topics. A meta-analysis may serve to combine findings from similar studies to help increase the power to detect statistical differences. Rather than base clinical practice on a single or few studies, meta-analysis allows for a pooled conclusion, thus adding to the strength and generalizability of the findings.

For example, a meta-analysis by the Cochrane Collaboration (2004) compared laparoscopic vs open appendectomy for suspected appendicitis.[9] The investigators identified many RCTs comparing these 2 procedures; however, none clearly delineated the benefits. The meta-analysis combined the results and found that the laparoscopic approach was associated with a lower wound infection rate, decreased incidence of intra-abdominal abscess, less postoperative pain, and shorter duration of hospital stay.

Another example is seen with the studies designed to compare the outcomes for breast cancer between lumpectomy plus radiation vs mastectomy. One of the first landmark studies on the this topic was the National Surgical Adjuvant Breast and Bowel Project (NSABP) Protocol B-06 trial, which showed that lumpectomy with radiation as compared with mastectomy had similar mortality rates and acceptable breast cancer recurrence rates.[10] However, because the sample size was relatively small (n=1529 at 12-year follow-up), it was important to demonstrate that these results were reproducible in other trials, particularly because dramatic changes in clinical practice were being undertaken. A meta-analysis by the Early Breast Cancer Trialists' Collaborative Group, looking at 36 trials, which included more than 29 000 women, helped to confirm the NSABP findings.[11] Their large analysis of all available studies confirmed that radiotherapy combined with surgery led to a local recurrence rate 3 times lower than the rate from surgery alone and there was no significant difference in mortality.[11] Meta-analysis works as a valuable method to combine all available surgical studies to validate findings from a single study or a few studies that may have potentially treatment-changing impact.

META-ANALYSIS: METHODS

Multistep Approach

Meta-analysis follows an approach similar to that used for research involving primary data collection.[6,7,12] The 5 steps include: (1) define research question, (2) establish study selection criteria, (3) perform literature search of topic, (4) abstract study features and outcomes, and (5) analyze data and present results.

The first step establishes a specific research question and a formal protocol detailing objectives and hypotheses. The research question should address the intervention, comparison group, and clinical outcomes of interest.[13] Within general surgery, an example would be to compare laparoscopic vs open inguinal hernia repair and select the outcomes to measure, such as recurrence rate and postoperative complications.

Downloaded from www.archsurg.com at National Institute of Hlth, on October 9, 2007
©2006 American Medical Association. All rights reserved.

The second step defines the inclusion and exclusion criteria for identifying eligible studies. Criteria should detail the type of study design (eg, RCT, observational cohort study), patients (eg, age, sex, presence of medical conditions), data publication (eg, abstract only or nonpublished data), language (eg, inclusion of non-English studies), and time period.[6,7] Additionally, many items specific to the surgical experience must be considered, including other procedure details (eg, primary hernia, method of repair, type of mesh), surgeon (eg, experience level), hospital (eg, teaching institution), or timeliness of study. For example, there may be discrepancy in treatment or clinical management between studies performed in the past 2 years vs those performed 20 years ago. The degree of criteria specificity can greatly impact overall results as broad inclusion criteria could increase heterogeneity among the surgical studies, while narrow inclusion criteria could lead to limited subgroup analysis.[14] For the inguinal hernia example, using too broad of inclusion criteria for the method of securing the mesh (interrupted vs running) in the open study arms could potentially cause too much heterogeneity when comparing outcomes between the laparoscopic vs open approach.

Third, a literature search is performed to obtain all relevant surgical studies that meet the inclusion criteria; the specific terms used to conduct the search should be reported. Common database search tools include MEDLINE, the Cochrane Library, Current Contents, and EMBASE. Using multiple databases is crucial to ensure that pertinent publications are not omitted; non-English–written articles should be included and translated when appropriate.[6,7,12] Scanning bibliographies of the retrieved articles and asking surgical experts in the field may identify additional publications.[7] Articles with duplicate or previously published data should be excluded. For example, an RCT may generate several publications on the same patient population so it is imperative not to double count them.

The fourth step involves abstraction of study features, characteristics of the patient population, and outcomes onto a standard data collection form.[12] The quality of studies should be evaluated with regard to randomization, double blinding, and explanation for dropouts and withdrawals, which addresses the issues of internal validity (minimization of bias) and external validity (ability to generalize results).[15] To maintain accuracy, 2 independent researchers should extract the data. The degree of agreement between the reviewers should be calculated (eg, κ statistic) and discrepancies, resolved.[12] Blinding researchers to the study authors and other identifying information may decrease the chance of bias.

Overall, particularly for surgery, few randomized articles on a given topic may be available and heterogeneity between those studies can be significant. As such, the majority of surgical meta-analyses are composed of relatively few articles, frequently fewer than 10. A recent review article by Dixon et al[16] looked at the quality of published meta-analysis within general surgery. They found 51 meta-analyses to review, and of these, 18 analyzed 10 or fewer studies. It is acceptable to perform meta-analysis on surgical topics with few available studies—even 3 is commonly performed—as long as the degree of heterogeneity is minimal.

Data Analysis and Presentation of Results

The fifth step of a meta-analysis includes data analysis and presentation of results. Selection of the specific type of analysis depends on whether the outcome variable is continuous (eg, postoperative length of hospital stay) or dichotomous (eg, postoperative complication event). For continuous end points, the mean difference between 2 groups (eg, control and treatment groups) is recorded.[12] Data must be translated to a common scale to allow for comparison. For example, a recent meta-analysis on bariatric surgery found that the majority of articles reported preoperative weight in kilograms, while some articles reported preoperative weight in pounds.[17] Transforming data into a common scale will allow for maximal inclusion of data.

If the end point is dichotomous (effect vs no effect), the OR or RR is calculated.[12] "Odds" is defined as the ratio of events to nonevents, and the OR is defined as the odds in 1 group (eg, treatment group) divided by the odds in a second group (eg, control group). An OR greater than 1 means that the event is more likely in the treatment group, and therefore, the treatment group is favored if the "event" is desirable (eg, 1-year survival). "Risk" is defined as the number of patients with an event divided by the total number of patients, and the risk ratio is defined as the risk in 1 group (eg, treatment group) divided by the risk in a second group (eg, control group). An RR less than 1 favors the treatment group if the "event" is not desirable (eg, reoperation). Relative risk is easier to interpret intuitively, but some study designs, like case control, prevent its calculation and OR should be used.

For RCTs, absolute measures, such as the risk difference (RD), also called the absolute risk reduction, and number of patients needed to treat (NNT) can be calculated. The RD is defined as the risk in the treatment group minus the risk in the control group, which quantifies the absolute change in risk due to the treatment. In general, a negative risk difference favors the treatment group. The NNT is defined as the inverse of the risk difference. The NNT provides the number of patients who need to be treated with the intervention to prevent 1 event. In the case where the risk difference is positive (does not favor the treatment group), the inverse will provide the number needed to harm (NNH).[18]

The meta-analysis technique for combining data also uses a weighted average of the results. Larger trials are given more weight since the results of smaller trials are more likely to be affected by chance.[6,12] Either the fixed-effects or random-effects model can be used for determining the overall effect. The fixed-effects model assumes that all studies are estimating the same common treatment effect; therefore, if each study were infinitely large, an identical treatment effect could be calculated. The random-effects model assumes that each study is estimating a different treatment effect and hence yields wider confidence intervals. A meta-analysis should ideally be analyzed using both the fixed-effects and random-effects models. If there is not a difference between the 2 models, then the studies are unlikely to have significant statistical heterogeneity. If there is a considerable difference between the 2 models, then the most conservative estimate should be reported, which is usually the random-effects model.[18]

When reporting a meta-analysis, the combined study effect (ie, difference between the study arms) is presented graphically along with the results of individual studies, as demonstrated by the forest plot example in **Figure 1**. In this example, hernia recurrence rate is compared between 2 procedures and the diamond-shaped figure represents the pooled results, with the midline corresponding to the calculated estimate for all studies and the horizontal spread representing the 95% confidence interval. Since the horizontal spread does not cross the vertical line (located at an OR of 1) in this forest plot, this demonstrates a statistically significant difference, with procedure X having a lower rate of hernia recurrences than procedure Y.

There are additional types of meta-analysis that can be performed. One approach is to run the analysis based on individual patient data.[19] Individual patient-level data involves the reanalysis of the actual raw data from all relevant RCTs. The researcher conducting the individual patient-level data study collects the data and reanalyzes it. Often, the data are double-

Downloaded from www.archsurg.com at National Institute of Hlth, on October 9, 2007
©2006 American Medical Association. All rights reserved.

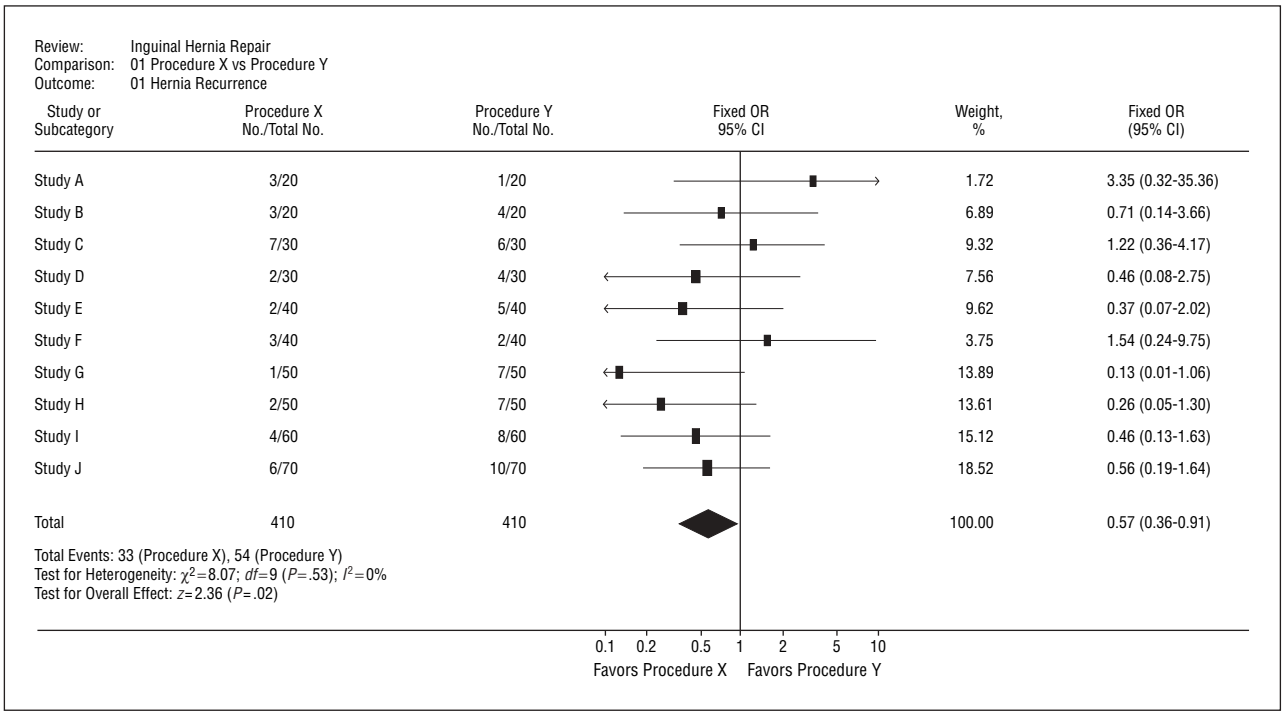

Figure 1. Illustration of a forest plot for a hypothetical meta-analysis comparing hernia recurrence incidence following hypothetical procedure X vs procedure Y for inguinal hernia repair. This hypothetical figure was created using Review Manager 4.2 software, which is used by the Cochrane Collaboration to perform formal Cochrane reviews.[18] Odds ratio (OR)=0.57 suggests that procedure X has lower odds of having recurrence (43% lower) than procedure Y. Test for heterogeneity reports χ^2=8.07, *df*=9, which usually is 1 less than the total number of studies (10−1=9), and corresponding *P* value of .53, which suggests no significant heterogeneity among the studies (in contrast, a *P* value <.10 implies possible heterogeneity). I^2=0% quantifies the degree of heterogeneity and negative values are converted into 0%; 0% corresponds with no observed heterogeneity, as is seen in this example. *z* Test of overall effect=2.36 with corresponding *P* value=.02 suggests that there was a significant effect difference between the 2 procedures. CI indicates confidence interval.

checked for accuracy and follow-up times are updated. This process involves obtaining the raw data from the original researchers who conducted each study, which may not always be possible. While this technique takes longer to perform and requires additional resources, it has several advantages including detailed subanalyses, flexibility of analysis, and standardization of outcome variables and follow-up. The Cochrane Collaboration has a section devoted to performing these types of studies, called the Individual Patient Data Meta-analysis Methods Group.[20] A recent example of this type of meta-analysis looked at repair of inguinal hernias and was called "Open Mesh Versus Non-Mesh Repair of Groin Hernia: Meta-Analysis of Randomised Trials Based on Individual Patient Data."[21] While this method requires a greater amount of resources, it lessens the degree of publication and selection bias, thus potentially resulting in more accurate results.

Another example is the cumulative meta-analysis, which involves repeating the meta-analysis as new study findings become available and allows for the accrual of data over time.[8] It can also retrospectively pinpoint the time when a treatment effect achieved statistical significance or identify if a treatment's effectiveness diminishes over time (ie, antibiotic resistance).[22] This can be particularly informative for surgical patients with cancer as adjuvant treatment evolves with time.

A surgical meta-analysis can also be performed using either observational or RCT data. Ideally, limiting the meta-analysis to only RCT data will produce results with a higher level of scientific evidence; however, there are often only a few of these studies available. Randomized data will be less likely to have significant selection bias or other confounding factors. Pooling nonrandomized data has many limitations that must be considered in the final assessment of the results. Furthermore, a general rule of thumb is that observational data should not be combined with randomized data for meta-analysis.

While a meta-analysis can produce an overall conclusion on a surgical topic with more power than looking at the individual studies, results must be interpreted with consideration of the study question, selection criteria, method of data collection, and statistical analysis.[6]

META-ANALYSIS: LIMITATIONS

The primary criticism of a meta-analysis is the presence of multiple types of bias, which is common within surgery. Pooling data from different sources unavoidably includes biases of the individual studies.[23,24] Moreover, despite the establishment of study selection criteria, authors may tend to incorporate studies that support their view, leading to selection bias.[6,7,23-26] There is also potential for bias in identification of studies because they are often selected by investigators familiar with the field who have individual opinions.[23] For example, an author who believes strongly in the benefits of the laparoscopic approach for colon cancer may unwittingly exclude studies that do not support his or her view.

Language bias may exist when literature searches fail to include foreign studies, because significant results are more likely to be published in English.[6,23] Studies with significant findings tend to be cited and published more frequently, and those with negative or nonsignificant findings are less likely to be published, resulting in possible citation bias and publication bias.[6,23] Since studies with significant results are more likely to be indexed in the literature database, database bias is another concern.[6,23] Studies that have not been published in traditional jour-

Downloaded from www.archsurg.com at National Institute of Hlth, on October 9, 2007
©2006 American Medical Association. All rights reserved.

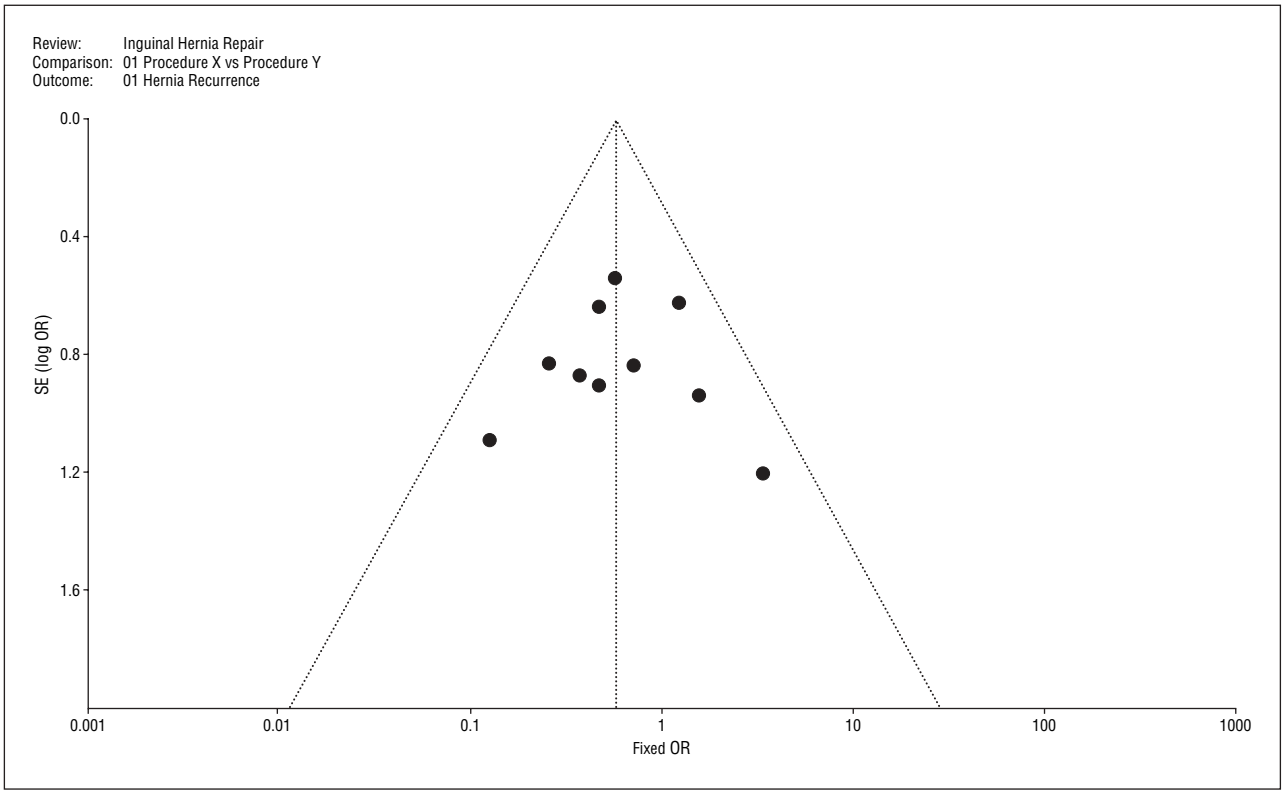

Figure 2. Illustration of a funnel plot for a hypothetical meta-analysis comparing hernia recurrence incidence following hypothetical procedure X vs procedure Y for inguinal hernia repair. The y-axis reflects individual study weights as the log of the effect estimate, SE (log OR). The x-axis represents the odds ratio (OR) for each study. The symmetry of the plot distribution suggests absence of publication bias. SE indicates standard error.

nals, like a dissertation or a chapter, are referred to as "fugitive" literature and are less likely to be identified through the traditional database search. Finally, multiple-publication bias can occur if several publications are generated from a multicenter trial or a large trial reporting on a variety of outcomes. If the same set of patients is included twice in the meta-analysis, the treatment effect can be overestimated.[23] These potent bias factors can affect the conclusions and must be considered during interpretation of the results.

To combat these sources of bias, several tools are available. First, a sensitivity analysis can help examine for bias by exploring the robustness of the findings under different assumptions.[23] Exclusion of studies based on specified criteria (eg, low quality, small sample size, or studies stopped early because of an interim analysis) should not significantly change the overall effect if the results of the meta-analysis are not significantly influenced by these studies. Second, the degree of study heterogeneity is another major limitation and the random-effects model should be used when appropriate and attempts made to identify potential causes (eg, clinical differences in the studies) of the heterogeneity.[18]

A third approach at measuring potential bias is the funnel plot, which is a scatterplot illustrating each study's effect with reference to their sample size (**Figure 2**). The underlying principle is that as the sample size of individual studies increases, the precision of the overall estimate improves. This is shown graphically because smaller studies would distribute widely while the spread of large studies should be narrow. The funnel plot shows a symmetrical inverted funnel if there is minimal or no bias, as demonstrated in Figure 2. By the same logic, the plot would be asymmetrical and skewed when bias exists.

The quality of a meta-analysis is dependent on the degree of heterogeneity between studies. The recent review by Sauerland and Seiler[5] details the importance of careful selection of studies while taking into account variations between treatments or patient populations that may exist. Particularly within surgery, where procedure techniques or patient selection may vary, it is imperative to identify the relevant differences and limit the statistical pooling of data to situations where heterogeneity is minimal. In fact, when too much heterogeneity between studies is apparent, data should not be pooled and definitive conclusions will not be able to be drawn until more studies become available.[25,26]

Also, the strength and precision of a meta-analysis is in question when the results contradict a large, well-performed RCT.[23,24] As such, results of any individual study or trial may be overlooked in place of the pooled results. However, it is arguable that findings falling outside the group mean are likely a product of chance and may not reflect the true effect difference, which provides the rationale for formally pooling similar studies. Even if a real difference exists in an individual trial, the results of the group will likely be the best overall estimate (also known as the Stein Paradox).[27]

Caution should be exercised when using subgroup analysis to make decisions on individual patients. Meta-analysis approximates the overall effect of a treatment in a wide range of subjects and thus subgroup analysis is

Downloaded from www.archsurg.com at National Institute of Hlth, on October 9, 2007
©2006 American Medical Association. All rights reserved.

susceptible to bias. Surgeons should consider the risks and comorbidities of the studied population in comparison with their own patients to help decide whether the findings are clinically applicable.

The recent review article of meta-analyses within general surgery by Dixon et al[16] found many inadequacies in the quality of these studies. Overall, the majority of the meta-analyses had major methodological flaws—median score of 3.3 on a scale from 1 to 7. Areas of weakness included errors in validity assessment, selection bias of patient populations, poor reporting of search strategies, and improper pooling of data. Further, they found that poor-quality meta-analyses tended to report a greater effect difference than the higher-quality ones. These results emphasize the importance of performing meta-analysis using rigorous and high-quality methods.

A recent article on the effectiveness of bariatric surgery highlights the strengths and limitations of these methods.[28] The study included both a meta-analysis and systematic review on the benefits and risks of bariatric surgery. Because there were few controlled studies comparing the effectiveness of surgery vs a control arm (ie, medical treatment for weight loss), pooling of this data was not possible. On the other hand, there were a handful of RCTs comparing outcomes for different types of bariatric surgery (ie, vertical banded gastroplasty vs gastric bypass) for which the data for several studies could be pooled and analyzed collectively.

The meta-analysis method has inherent limitations; however, attempts by researchers to decrease bias, minimize study heterogeneity, and carefully interpret results will improve the likelihood of generating valid and accepted conclusions.

CONCLUSIONS

This overview aimed to clarify some of the concepts behind meta-analysis as well as review the basics for those more familiar with the method. Like primary research, meta-analysis involves a stepwise approach to arrive at statistically justifiable conclusions. It has the potential to provide an accurate appraisal of the literature along with quantitative summation and may objectively resolve controversies. The greatest challenge in conducting a meta-analysis on a surgical topic is often the lack of available data on the subject, because there are few high-quality, published studies with an acceptable degree of heterogeneity.

Since the number of meta-analyses published in recent years has increased substantially, it is imperative that surgeons understand the strengths and weaknesses of this method.[7,29] Furthermore, surgeons must also have the ability to critically judge the findings since the results of a meta-analysis have the potential to influence clinical practice.

Accepted for Publication: August 10, 2005.
Correspondence: Melinda A. Maggard, MD, MSHS, Department of Surgery, David Geffen School of Medicine at UCLA, 72-215 CHS, 10833 Le Conte Ave, Los Angeles, CA 90095 (mmaggard@mednet.ucla.edu).
Author Contributions: *Study concept and design:* Ng, McGory, and Ko. *Acquisition of data:* Ng and McGory. *Analysis and interpretation of data:* Ng, McGory, Ko, and Maggard. *Drafting of the manuscript:* Ng, McGory, and Maggard. *Critical revision of the manuscript for important intellectual content*: Ko and Maggard. *Statistical analysis:* McGory, Ko, and Maggard. *Administrative, technical, and material support:* Ng. *Study supervision:* Ko and Maggard.
Financial Disclosure: None reported.

REFERENCES

1. Glass GV. Primary, secondary, and meta-analysis of research. *Educ Res.* 1976; 5:3-8.
2. Pearson K. Report on certain enteric fever inoculation statistics. *BMJ.* 1904;3:1243-1246.
3. Lasagna L, Mosteller F, Von Felsinger JM, Beecher HK. A study of the placebo response. *Am J Med.* 1954;16:770-779.
4. Cochrane AL. *Effectiveness and Efficiency: Random Reflections on Health Services.* London, England: Nuffield Provincial Hospitals Trust; 1972.
5. Sauerland S, Seiler CM. Role of systematic reviews and meta-analysis in evidence-based medicine. *World J Surg.* 2005;29:582-587.
6. Wolf FM. Introduction to systematic reviews and meta-analysis. http://depts.washington.edu/k30/Meta-analysis/Meta-analysis%20clinical%20research%200603_files/frame.htm. Accessed November 11, 2004.
7. Kelly GA. Meta-analysis: an introduction. http://www.pitt.edu/~super1/lecture/lec3221/index.htm. Accessed November 11, 2004.
8. Egger M, Smith GD. Potentials and promise. *BMJ.* 1997;315:1371-1374.
9. Sauerland S, Lefering R, Neugebauer E. Laparoscopic versus open surgery for suspected appendicitis. *Cochrane Database Syst Rev.* 2004;(4):CD001546.
10. Fisher B, Anderson S, Redmond CK, Wolmark N, Wickerham DL, Cronin WM. Reanalysis and results after 12 years of follow-up in a randomized clinical trial comparing total mastectomy with lumpectomy with or without irradiation in the treatment of breast cancer. *N Engl J Med.* 1995;333:1456-1461.
11. Early Breast Cancer Trialists' Collaborative Group. Effects of radiotherapy and surgery in early breast cancer—an overview of the randomized trials [published correction appears in *N Engl J Med.* 1996;334:1003]. *N Engl J Med.* 1995; 333:1444-1455.
12. Egger M, Smith GD, Phillips AN. Meta-analysis: principles and procedures. *BMJ.* 1997;315:1533-1537.
13. Meade MO. Selecting and appraising studies for a systematic review. *Ann Intern Med.* 1997;127:531-537.
14. Counsell C. Formulating questions and locating primary studies for inclusion in systematic reviews *Ann Intern Med.* 1997;127:380-387.
15. Moher D, Pham B, Jones A, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet.* 1998; 352:609-613.
16. Dixon E, Hameed M, Sutherland F, Cook DJ, Doig C. Evaluating meta-analyses in the general surgical literature: a critical appraisal. *Ann Surg.* 2005;241:450-459.
17. Buchwald H, Avidor Y, Braunwald E, et al. Bariatric surgery: a systematic review and meta-analysis. *JAMA.* 2004;292:1724-1737.
18. The Cochrane Collaboration Web site. http://www.cochrane.org. Accessed November 11, 2004.
19. Stewart LA, Clarke M. Practical methodology of meta-analyses (overviews) using updated individual patient data: Cochrane Working Group. *Stat Med.* 1995; 14:2057-2079.
20. Cochrane Collaboration Individual Patient Data Meta-analysis Methods Group. http://212.219.75.236/ukcccr/ipd/home.asp. Accessed July 31, 2005.
21. Grant AM; EU Hernia Trialists Collaboration. Open mesh versus non-mesh repair of groin hernia: meta-analysis of randomised trials based on individual patient data [published correction appears in *Hernia.* 2002;6:204]. *Hernia.* 2002; 6(3):130-136.
22. Chang DC, Wilson SE. Meta-analysis of the clinical outcome of carbapenem monotherapy in the adjunctive treatment of intra-abdominal infections. *Am J Surg.* 1997; 174:284-290.
23. Egger M, Smith GD. Bias in location and selection of studies. *BMJ.* 1998;316:61-66.
24. LeLorier J, Grégoire G, Benhaddad A, Lapierre J, Derderian F. Discrepancies between meta-analyses and subsequent large randomized, controlled trials. *N Engl J Med.* 1997;337:536-542.
25. Bailar JC III. The practice of meta-analysis. *J Clin Epidemiol.* 1995;48:149-157.
26. Bailar JC. The promise and problems of meta-analysis. *N Engl J Med.* 1997;337: 559-561.
27. Efron B, Morris G. Stein's Paradox in statistics. *Sci Am.* 1977;236:119-127.
28. Maggard MA, Shugarman LR, Suttorp M, et al. Meta-analysis: surgical treatment of obesity. *Ann Intern Med.* 2005;142:547-559.
29. Davey Smith G, Egger M. Meta-analysis: unresolved issues and future developments. *BMJ.* 1998;316:221-225.

Downloaded from www.archsurg.com at National Institute of Hlth, on October 9, 2007
©2006 American Medical Association. All rights reserved.