

Role of systematic reviews and meta-analysis in evidence-based clinical practice

Erin R. McNamara, Charles D. Scales Jr

Division of Urologic Surgery, Department of Surgery Duke University Medical Center, Durham, NC, USA

ABSTRACT

Introduction: Systematic reviews and meta-analyses of well-designed and executed randomized controlled trials have the potential to provide the highest levels of evidence to support diagnostic and therapeutic interventions in urology.

Materials and Methods: The role of systematic reviews in the urological literature is described. A three-step appraisal of the validity, magnitude and applicability of results will permit an evidence-based approach to incorporating findings of systematic reviews and meta-analyses into practice.

Results: The validity of systematic reviews depends on a focused clinical question that generates specific inclusion and exclusion criteria for identifying studies through an exhaustive literature search. The primary studies must be of high methodological quality and assessments should be reproducible. Informed consumers of the urological literature should be aware of the consistency of results between trials in a review, as well as the magnitude and precision of the best estimate of the treatment effects. When making decisions about implementing the results, urologists should consider all patient-important outcomes, the overall quality of the evidence and the balance between benefits, potential harms and costs.

Conclusion: This framework will lead to a more evidence-based application of systematic reviews within the urological literature. Ideally, utilization of an evidence-based approach to systematic reviews will improve the quality of urological patient care.

Key words: Evidence-based medicine, systematic review, meta-analysis, randomized controlled trials

INTRODUCTION

One of the fundamental principles of evidence-based practice is that a hierarchy of reliable evidence exists, based on design elements that minimize the effect of bias on the results.^[1-3] Well-done randomized controlled trials provide high levels of evidence for questions of therapy. When multiple randomized controlled trials exist which address the same clinical question a systematic review and meta-analysis of the trial results have the potential to provide the highest level of evidence. The purpose of this manuscript is to

introduce the reader to a method for critical appraisal of a systematic review and meta-analysis. Using an example from the urological literature, the reader will be guided through a three-step approach to assessing and applying the results of a systematic review and meta-analysis to clinical practice.

CASE SCENARIO

One day you are in the surgeons' lounge talking with a general surgery colleague between cases. He has recently returned from a surgical conference, and asks your opinion about suprapubic catheterization (SPC) as compared with transurethral catheterization (TUC) in patients undergoing surgery. He says that another general surgeon at the conference always places suprapubic catheters at the time of laparotomy, because there is high level evidence that it decreases urinary tract infections and patient discomfort. At the hospital where both of you practice, TUC has been the customary practice. Intrigued, you tell him that you are not aware of high level evidence about this question, but you resolve to look into this issue after your cases are done for the day.

LITERATURE SEARCH

Having recently read an article about searching the medical literature,^[4] you decide to first see if a preappraised source

For correspondence: Dr. Charles Scales Jr, DUMC Box 2922, Division of Urologic Surgery, Duke University Medical Center, Durham, NC-27710, USA
E-mail: chuck.scales@duke.edu

Access this article online	
Quick Response Code:	Website: www.indianjurol.com
	DOI: 10.4103/0970-1591.91445

of evidence is available on this topic. You recall that ideally several well-done randomized controlled trials addressing the same question that yield similar results comprise the highest level of evidence, and are frequently summarized in a systematic review and meta-analysis. The Cochrane Collaboration is a non-profit organization which assembles high-quality systematic reviews of healthcare interventions. These cochrane reviews (<http://www.cochrane.org/cochrane-reviews>) are available online, and you access the website to begin your literature search. Typing in the term 'urinary catheterization' and limiting the results to reviews, you identify 40 studies. While you would prefer to scan fewer titles, the cochrane review search engine sorts according to relevance. The fifth review on your list is titled 'Urinary catheter policies for short term bladder drainage in adults' (search date May 10, 2010).^[5] Reviewing the abstract, you note that this study reviewed trials comparing routes of insertion for short term (up to 14 days) urinary catheterization in adults. This appears to directly address your question, so you download the full study for review. The study identified and included 14 randomized controlled trials which compare suprapubic to urethral catheterization for bladder drainage after various surgical procedures in adults. In addition, the review identified and included three studies of intermittent versus indwelling urethral catheterization as a separate comparison.

SYSTEMATIC REVIEWS

The biomedical literature is replete with articles reviewing specific disorders or interventions, but *systematic reviews* hold a special place within this group. Traditional, or narrative reviews, typically provide an overview of a condition and address a number of questions surrounding the diagnosis, prognosis or management of that condition, and are heavily weighted with expert opinion.^[1] In contrast, a systematic review addresses a focused clinical question while using specific methods to minimize introduction of bias during the review process. Often times a systematic review will also include a meta-analysis, which is a quantitative calculation of the overall treatment effect from the data presented in each trial.^[1] The key difference to remember is that when done well, a systematic review (and meta-analysis, if included) contains specific methodological safeguards against introduction of bias. In order to critically appraise a systematic review, we will sequentially apply three questions [Table 1]: 1) Are the results valid? 2) What are the results? and 3) How can I apply the results to patient care?

ARE THE RESULTS VALID?

Standards for conduct and reporting of systematic reviews and meta-analyses exist.^[6] Among these standards, a few key criteria should always be assessed when appraising a systematic review. These include the eligibility criteria

Table 1: Criteria for assessing a systematic review. Adapted from *Users' Guides to the Medical Literature*.^[1]

Are the results of the study valid?
Did the review include explicit and appropriate eligibility criteria?
Was biased selection and reporting of studies unlikely?
Were the primary studies of high methodological quality?
Were assessments of studies reproducible?
What are the results?
Were the results similar from study to study?
What are the overall results of the review?
How precise were the results?
How can I apply the results to the care of my patient?
Were all patient-important outcomes considered?
Are any postulated subgroup effects credible?
What is the overall quality of the evidence?
Are the benefits worth the costs and potential risks?

for study inclusion, potential for biased selection or reporting of the studies, an assessment of methodological quality, and whether or not assessments of the studies were reproducible.

Did the review include explicit and appropriate eligibility criteria?

A systematic review should build around a focused clinical question, and then specify appropriate inclusion and exclusion criteria for selecting the underlying studies which will comprise the systematic review. It is important to understand the patient population, treatment, and outcomes so that the urologist can decide whether pooling of the studies is appropriate. In addition, the eligibility criteria should include some sort of methodological criteria. For example, in a systematic review of α -blocker therapy for benign prostatic hyperplasia symptoms, it would be appropriate to specify that only randomized controlled trials be included, since they offer the highest level of evidence.

For our clinical scenario, the review by Niel-Wiese *et al.*, explicitly states that only randomized controlled trials and quasi-randomized trials comparing alternative approaches to short term catheterization in hospitalized adults will be included in the review.^[5] The patient population included adults with general medical problems, acute urinary retention, acute illness and those undergoing surgery. The interventions included urethral catheterization, SPC, and intermittent catheterization. A number of outcome measures, including patient comfort, quality of life, complications and costs could be included. After review of the methods section, we can conclude that the authors appropriately defined trial inclusion criteria, although whether general medical patients and surgical patients are similar enough to analyze together may require further assessment based on the results of the underlying studies.

Was biased selection and reporting of studies unlikely?

In order for systematic reviews to achieve an accurate representation of the treatment effect for a given intervention, it is critical that all of the available eligible studies be identified and included. Authors should specifically report the search strategy for the systematic review. Good search strategies typically include bibliographic databases (i.e., MEDLINE, EMBASE), abstracts published at scientific meetings, hand searches of the reference lists of published studies, and registries of ongoing research.^[1,7] In particular, reporting bias may occur with negative studies that appear, if published at all, in less well-known journals.^[8,9] Failure to include negative studies may lead to inaccurate estimates of treatment effects.^[10,11]

The authors, in our scenario, report an exhaustive search strategy, including searches of MEDLINE, CINAHL and the Cochrane Central Register of Controlled Trials (CENTRAL). In addition, reference lists and conference proceedings were hand searched. The authors specify that language restrictions were not imposed, so trials reported in languages other than English were not excluded. Overall, we can conclude that biased selection and reporting of studies was unlikely.

Were primary studies of high methodological quality?

It is important that the quality of the primary studies be high enough to minimize the potential for biased estimates of treatment effects, or at least that the authors of the systematic review assess and report potential weaknesses in study methods. In reviews of therapeutic interventions, several key elements of randomized controlled trials should be assessed, including randomization, allocation concealment, appropriate blinding, analysis according to the intention-to-treat principle, and appropriate follow-up.^[2,12,13] Empiric evidence suggests that biased results are associated with failure to report these key methodological elements.^[14-16] In the review on catheterization routes, the authors explicitly describe their quality assessment, which included randomization method, allocation concealment, description of dropout and withdrawals, intention to treat analysis and blinding. Overall, the authors were concerned about quality limitations in the primary studies.^[5] Specifically, details of randomization and allocation concealment were either suboptimal or not reported, and high rates of dropout were observed in some trials. As critical appraisers of this systematic review, it appears that the authors appropriately assessed the primary studies, but the results of those studies were subject to potential bias due to suboptimal methods. Therefore, we have to keep this limitation in mind when reading the results of the review.

Were assessments of studies reproducible?

In the process of conducting a systematic review, several decisions occur, including study inclusion, quality assessment, and data abstraction. Each of these decisions may include subjective and objective components, which

introduce the opportunity for both random errors and bias. One way to minimize the opportunity for errors is to use at least two independent reviewers at each stage of the review. Statistical tests also exist that can measure the degree of agreement beyond chance. The Cochrane review of catheterization routes used two independent reviewers to select studies, assess quality of the primary studies, and extract data.^[5] Disagreements were resolved by consensus or an independent third party. Statistical measures of agreement were not reported, but overall the authors appear to have made reasonable efforts toward reproducible study assessments.

WHAT ARE THE RESULTS?**Were the results similar from study to study?**

Having assessed the validity of the systematic review, we must now turn to the results of the study. First, we must decide whether it make sense to combine results from several different studies across different patients and care settings. In the case of the systematic review and meta-analysis of catheterization route, there were 14 trials including patients with gynecological surgery, laparotomy, and other general surgical procedures. Does it make sense to combine these populations? Do we expect the treatment effect of urethral or SPC to be similar among these trials? Intuitively, we have identified the concept of heterogeneity—that is, differences between trials that may impair our ability to reasonably combine results.

How do we assess for heterogeneity among primary studies when evaluating a systematic review and meta-analysis? Four primary methods exist.^[1] First, we compare the point estimates of treatment effect for each study. If the point estimates of treatment effect vary widely from study to study, this causes concern that pooling results may not be appropriate. Alternately, if the point estimates are similar, then it appears to make sense to combine these data.

Second, we inspect the confidence intervals around the point estimates of the primary studies. In this case, we would like to see overlapping confidence intervals between studies, which reassures us that heterogeneity is low. On the other hand, if few of the confidence intervals overlap, reasons for the wide variation in results should be sought.

Two other methods for evaluating heterogeneity between studies utilize statistical tests. A hypothesis testing approach generates a *P*-value, which when less than 0.05 typically indicates that heterogeneity between study results exists beyond that expected by chance. Alternately, the *I*² test quantifies the amount of variability due to differences other than chance. Further details of these tests are beyond the scope of this review, but may be found in the Users' Guides to the Medical Literature.^[1] In the review of catheterization methods, the authors specifically note that they considered

the possibility of clinical heterogeneity in their analysis. However, only one outcome (asymptomatic bacteriuria) had a significant number of primary trials; in this case, the authors did note heterogeneity among studies, but report that subgroup analyses generally favored SPC.^[5]

What are the overall results of the study?

In the process of conducting a systematic review, investigators are frequently able to perform a meta-analysis, which can produce a single best estimate of the effect of an intervention. For dichotomous outcomes (i.e., alive or dead), the results are typically presented as relative risks or odds ratios. The contribution of each study to the overall estimate is typically weighted by the inverse of its variance. Thus, larger studies with less variance contribute more to the overall treatment effect estimate than smaller, high variance studies.^[7,17]

The authors present results from several endpoints, not all of which were included in every trial. Thus, the number of primary studies for each outcome varies. Asymptomatic bacteriuria was the most common outcome reported for the comparison of urethral versus SPC. In this comparison, bacteriuria was more common following urethral catheterization. The relative risk for asymptomatic bacteriuria following urethral catheterization was 2.60 times that of SPC. The authors report an absolute risk difference of 0.24, which translates into four patients requiring SPC to avoid one case of asymptomatic bacteriuria from urethral catheterization. SPC also resulted in less need for recatheterization, and less pain or discomfort for patients (data from 8 and 6 trials, respectively).^[5]

How precise were the results?

Even with the power conferred by combining the results of several randomized controlled trials in a meta-analysis, the end product remains only an estimate of the treatment effect. In order to quantify this uncertainty, investigators calculate a confidence interval (typically 95%) surrounding the estimate of the treatment effect. For the asymptomatic bacteriuria outcome, the 95% confidence interval is 2.12-3.18. This suggests that there is a 95% probability that the true relative risk of asymptomatic bacteriuria from urethral catheterization is between 2.12 and 3.18 times that of SPC.

HOW CAN I APPLY THE RESULTS TO THE CARE OF MY PATIENT?

Were all patient-important outcomes considered?

A number of different outcomes can result from a given intervention. For example, in the case of urethral or SPC, patient-important outcomes could include symptomatic bacteriuria, need for recatheterization, duration of catheterization, and catheter obstruction. In addition, quality-of-life measures such as pain or discomfort could be important to assess. Finally, adverse effects of interventions may be important, although

these are frequently underreported in randomized controlled trials published in the urology literature.^[18]

The Cochrane Review of short-term catheterization routes includes several different outcomes, including pain/discomfort. However, data regarding several patient-important outcomes were only present in 1 or 2 primary studies, so it is difficult to summarize these treatment effects. In addition, the primary outcome is asymptomatic bacteriuria. It is unlikely that asymptomatic bacteriuria, likely representing colonization of catheter biofilm, is a patient-important outcome (as opposed to a true urinary tract infection). Therefore, the review, while of high methodological quality, may not be able to inform our decision-making because of the outcomes of the primary studies.

Are any postulated sub-group effects credible?

Interpretation of sub-group effects carries a much higher risk of identifying results that appear statistically significant by chance (Type I, or false positive error). Increasing the number of subgroup analyses increases the risk of erroneously detecting a treatment effect. In the review of catheterization routes, the authors prespecify the subgroup comparisons. Subgroup comparisons also appear to be consistent across studies for the primary outcome of asymptomatic bacteriuria. The authors did test a relatively large number of subgroups, which increases the risk of a Type I error. However, each of the subgroup analyses favored the suprapubic route of catheterization, so we are less concerned than if there were apparent differences favoring one subgroup but not another.

What is the overall quality of the evidence?

Before applying the results of systematic reviews to patient care, we should consider the overall quality of the evidence for each patient-important outcome. For randomized controlled trials that are well designed and executed, the quality of evidence may be considered high. However, if the trials do not contain important design features to safeguard against bias, then the risk of bias is higher, and the quality of evidence must be downgraded.

In the review of catheterization routes, the authors noted that description of the randomization and allocation concealment processes were either absent, or when reported, were suboptimal. In addition, other design features between trials, such as use of perioperative antibiotics, were not standardized, which also raises concerns of interpretation. Finally, we note that many of the primary studies were conducted over a decade ago, raising concern about the generalizability of the findings to current surgical practice.

Are the benefits worth the costs and potential risks?

Ideally, a systematic review would be able to report cost measures as an outcome, as well as the potential adverse effects of an intervention. Weighing the clinical importance

of the outcome is critical to determining whether a potential intervention is worth the costs and potential risks. In determining catheterization route based on the results of the Cochrane Review, we note that the primary benefit seems to be reducing the rate of asymptomatic bacteriuria, which is of uncertain clinical importance. Other results, such as reduced need for recatheterization and reduced patient discomfort are important, but we know little about potential harms of the intervention, as noted by the authors. The reporting of adverse events in the primary studies of catheterization route is less than optimal, as noted by the authors. Finally, we have no cost data regarding the two interventions. Thus, it is hard to decisively embrace a conversion to SPC given the uncertain importance of the benefit and the lack of data on potential risks and costs.

RESOLUTION OF CASE SCENARIO

Overall, you conclude that the systematic review is methodologically strong, but the quality of the results is somewhat limited by the weaknesses of the underlying primary studies. The potential risks of suprapubic catheter insertion concern you, as does the unclear clinical importance of asymptomatic bacteriuria. Finally, you note that most of the primary studies are somewhat dated, and you are unsure if the results are applicable in your current practice, which has evolved rapidly over the past few years. You encounter your surgical colleague in clinic the next day, and outline your concerns about the 'high quality' evidence from a systematic review. Both of you decide that at this time you do not want to change your practice of urethral catheterization.

CONCLUSION

Critical appraisal of systematic reviews and meta-analyses can provide important insight into the utility of the results to patient care. A three-step approach addressing the validity, magnitude and applicability of the results will allow the reader to rapidly assess a systematic review. Understanding the quality of the primary studies is especially important given limitations in the quality of trials in the urological literature.^[12,18-21] This approach will improve urologists' ability to utilize an evidence-based clinical practice, and increase the quality of patient care.

ACKNOWLEDGMENT

The content of this review is based in part on concepts introduced in the *Users' Guides to the Medical Literature*^[1] and the *User's Guide to the Urological Literature*.^[4,7]

REFERENCES

- Guyatt G, Rennie D, Meade MO, Cook DJ. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. 2nd ed. New York: The McGraw-Hill Companies Inc.; 2008.
- Scales CD Jr, Preminger GM, Keitz SA, Dahm P. Evidence based clinical practice: A primer for urologists. *J Urol* 2007;178:775-82.
- Brighton B, Bhandari M, Tornetta P 3rd, Felson DT. Hierarchy of evidence: From case reports to randomized controlled trials. *Clin Orthop Relat Res* 2003;413:19-24.
- Krupski TL, Dahm P, Fesperman SF, and Schardt CM. User's guide to the urologic literature: How to perform a literature search. *J Urol* 2008;179:1264-70
- Niel-Weise BS, van den Broek PJ. Urinary catheter policies for short-term bladder drainage in adults. *Cochrane Database Syst Rev* 2005, Issue 3, Art. No.: CD004203.
- Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: The QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet* 1999;354:1896-900.
- Tseng TY, Dahm P, Poolman RW, Preminger GM, Canales BJ, Montori VM. How to use a systematic literature review and meta-analysis. *J Urol* 2008;180:1249-56.
- Smith WA, Cancel QV, Tseng TY, Sultan S, Vieweg J, Dahm P. Factors associated with the full publication of studies presented in abstract form at the annual meeting of the American Urological Association. *J Urol* 2007;177:1084-8.
- Dickersin K, Min YI, Meinert CL. Factors influencing publication of research results. Follow-up of applications submitted to two institutional review boards. *JAMA* 1992;267:374-8.
- Dickersin, K. The existence of publication bias and risk factors for its occurrence. *JAMA* 1990;263:1385-9.
- Cook DJ, Guyatt GH, Ryan G, Clifton J, Buckingham L, Willan A, *et al*. Should unpublished data be included in meta-analyses? Current convictions and controversies. *JAMA* 1993;269:2749-53.
- Scales CD Jr, Norris RD, Keitz SA, Peterson BL, Preminger GM, Vieweg J, *et al*. A critical assessment of the quality of reporting of randomized, controlled trials in the urology literature. *J Urol* 2007;177:1090-4.
- Bajammal S, Dahm P, Scarpero HM, Orovan W Bhandari M. How to use an article about therapy. *J Urol* 2008;180:1904-11.
- Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408-12.
- Noseworthy JH, Ebers GC, Vandervoort MK, Farquhar RE, Yetisir E, Roberts R. The impact of blinding on the results of a randomized, placebo-controlled multiple sclerosis clinical trial. *Neurology* 1994;44:16-20.
- Lijmer JG, Mol BW, Heisterkamp S, Bossel GJ, Prins MH, van der Meulen JH, *et al*. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999;282:1061-6.
- Oxman AD, Cook DJ, Guyatt GH. Users' guides to the medical literature. VI. How to use an overview. Evidence-Based Medicine Working Group. *JAMA* 1994;272:1367-71.
- Breau RH, Gaboury I, Scales CD Jr, Fesperman SF, Watterson JD, Dahm P. Reporting of harm in randomized controlled trials published in the urological literature. *J Urol* 2010;183:1693-7.
- Scales CD Jr, Norris RD, Peterson BL, Preminger GM, Dahm P. Clinical research and statistical methods in the urology literature. *J Urol* 2005;174:1374-9.
- Scales CD Jr, Norris RD, Preminger GM, Vieweg J, Peterson BL, Dahm P. Evaluating the evidence: Statistical methods in randomized controlled trials in the urological literature. *J Urol* 2008;180:1463-7.
- Breau RH, Carnat TA, Gaboury I. Inadequate statistical power of negative clinical trials in urological literature. *J Urol* 2006;176:263-6.

How to cite this article: McNamara ER, Scales CD. Role of systematic reviews and meta-analysis in evidence-based clinical practice. *Indian J Urol* 2011;27:520-4.

Source of Support: Nil, **Conflict of Interest:** None declared.