How to appraise the effectiveness of treatment

Suzanne B. Stewart, Phillip Dahm, Charles D. Scales Jr

Division of Urology, Department of Surgery, Duke University Medical Center, Durham, North Carolina; Department of Urology, College of Medicine, University of Florida, Gainesville, Florida, USA

ABSTRACT

Introduction: Advances in treatment and disease prevention occur frequently in urology. Urologists must identify a framework within which to evaluate these therapeutic innovations.

Materials and Methods: The evidence-based approach to critical appraisal is described using an example from the urological literature. A three-part assessment of the trial validity, treatment effect, and applicability of results will permit the urologist to critically incorporate medical and surgical advances into practice.

Results: Validity of clinical trials hinges upon balancing patient prognosis at the initiation, execution, and conclusion of the trial. Readers should be aware of not only the magnitude of the estimated treatment effect, but also its precision. Finally, urologists should consider all patient-important outcomes as well as the balance of potential benefits, harms, and costs, and patient values and preferences when making treatment decisions.

Conclusion: Use of this framework for critical appraisal will lead to a more evidence-based application of new therapies for patients. Incorporation of a more evidence-based practice within urology will lead to an increase in the quality of patient care.

Key words: Dutasteride, evidence-based medicine, randomized controlled trial

INTRODUCTION

Therapeutic innovations occur frequently in urology, and urologists seeking to provide up-to-date care for their patients should employ these new interventions in a manner which appropriately balances the risks and benefits of the therapy. Ideally, evidence-based clinical practice (EBCP) will form the basis for these decisions. EBCP has been defined as the “conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients.”[1] EBCP recognizes that a hierarchy of evidence exists, and that clinicians should incorporate clinical expertise and patient values and preferences along with best evidence to arrive at treatment decisions.[2]

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

Ideally, a systematic review and meta-analysis of several randomized controlled trials (RCTs) will exist to guide treatment decisions. However, RCTs comprise a very small proportion of the urologic literature,[3] which inhibits meta-analysis. Since clinical research methods and trial reporting in the urology literature are frequently suboptimal,[3,4] urologists should understand how to critically appraise RCTs prior to applying the results to patient care. This review describes the skills for critical appraisal of RCTs, using an example from the urologic literature.

CLINICAL SCENARIO

A 50-year-old male returns to your clinic for follow-up 3 months after a negative prostate biopsy. His prebiopsy PSA was 4.3 ng/mL, and his prostate exam reveals an approximately 30 gram prostate gland without nodularity or induration. He has minimal voiding symptoms. Despite the negative prostate biopsy, the patient is particularly concerned about prostate cancer, and asks you what he can do to reduce his risk of prostate cancer in the future.

The patient mentions that he recently heard a news report about a trial of medication for enlarged prostates that may also reduce the risk of prostate cancer. You recall that the results of an RCT to evaluate the effect of dutasteride on prostate cancer risk were just released, but you have not yet appraised the trial report.[6] You tell the patient that
you too have heard the media reports, but would like to review the trial publication before making a treatment recommendation for him. The patient is very interested in your appraisal of the evidence, and sets an appointment to see you in two weeks.

THE LITERATURE SEARCH

Literature searches are ideally guided by a clinical question, in the PICOT format.[5] PICOT stands for Patient, Intervention, Comparison, Outcome, and Type of study. In the present scenario, one clinical question we could formulate is “in men without prostate cancer, do 5-alpha reductase inhibitors or placebo decrease the risk of prostate cancer?” In the absence of a meta-analysis, we would like to find at least one RCT. Using Pubmed, you enter the terms “prostate cancer,” “5-alpha reductase inhibitor,” and “cancer prevention.” Combining these search terms yields 129 articles, which is too many to peruse given your limited time. You decide to limit these results using the term “randomized controlled trial,” which results in 33 articles (date of search: April 11, 2010). With a quick review of the results, you find a RCT entitled “Effect of Dutasteride on the Risk of Prostate Cancer”[6] (REDUCE trial) and review the manuscript.

CRITICAL ASSESSMENT

Having located the study of interest, we must decide how to assess the findings. An evidence-based approach emphasizes sequentially asking three interdependent questions: (1) Are the results valid? (2) What are the results? (3) How can I apply these results to care of an individual patient? In the following sections, we will examine the REDUCE trial using the lens of this three question approach.

ARE THE RESULTS VALID?

Did experimental and control groups begin the study with similar prognosis?

Well-executed RCTs contain several design features to minimize potential bias in the results. Here, bias specifically means systematic deviation from the truth, as opposed to random error.[2] It is important to understand that bias may be introduced into study results at the beginning, middle, or end of a trial. Therefore, important methodological safeguards, which minimize bias should be reported for any RCT. At the beginning of an RCT, subjects in the experimental and control groups should have a similar prognosis. In order to minimize prognostic differences, patients should be randomized, the randomization process should be concealed, and a balance of known prognostic factors should exist between members of each group in the trial.

Were patients randomized?

The purpose of randomization is to balance both known and unknown prognostic factors between control and experimental groups. When successful, randomization assures us that the only prognostic difference between experimental and control groups is the treatment under investigation, and thus, any observed effect of therapy is due to that treatment.

In this case, the authors describe the trial as randomized, and refer the reader to a prior publication.[8] After a one month placebo run-in period, subjects were randomly assigned in a 1:1 ratio to active medication (dutasteride) or placebo. Randomization was stratified by center in this multicenter trial, which assures that each center will have similar numbers of patients in the control and experimental groups. If randomization were not stratified by center, then assignment to treatment or control groups could become unbalanced at a specific trial site, even across the entire trial subjects are evenly distributed across treatment and control groups. If enrollment from a particular site is somehow associated with an unknown prognostic factor, then not balancing treatment and control groups at each site could introduce bias into the results. Randomization in large, multicenter controlled trials is almost always performed by computer algorithm.

Was randomization concealed?

Concealment of randomization is another important concept in assuring that patients entering a trial share a similar prognosis. Essentially, concealment means that study personnel who enroll patients cannot predict the group assignment (experimental or control) of the next subject. Awareness of the allocation for the next subject may consciously or unconsciously influence an investigators decision to enroll a particular patient in the trial. Lack of concealment, or poor reporting of concealment, has been empirically associated with bias in RCTs.[9] In the REDUCE trial, whether and how randomization was concealed is not explicitly reported. In a large, multicenter trial, concealment is frequently accomplished by the use of a centralized randomization center.

Were known prognostic factors balanced between experimental and control groups?

Unfortunately, randomization is not perfect, and sometimes deviations from a balanced allocation can occur by chance, particularly if the sample size is small. For example, if only 12 patients were enrolled in a trial to reduce the incidence of prostate cancer, it would be easy to imagine that all 6 patients in the trial with a family history of the disease could by chance be allocated to the treatment arm, potentially creating a biased measurement of the treatment effect. However, if the trial enrolled 1200 patients, the chance of 600 patients with a family history of prostate cancer being allocated to a single trial arm is very low. While
we cannot assess the balance of unknown prognostic factors, investigators should report a table of known prognostic factors for comparison. Randomization should result in a balance of these measured prognostic factors. In the REDUCE trial, Table 1 summarizes relevant demographic and clinical variables for the dutasteride and placebo arms. By inspecting this table, we can see that there are no major differences between the treatment and control groups, which reassures us that randomization was in fact successful in balancing known (and unknown) prognostic factors.

**Was prognostic balance maintained as the study progressed?**

Were important groups blinded to treatment allocation, to the extent feasible?

When a particular group within a trial, such as the patients or the clinicians, is unaware of the treatment allocation, that group is referred to as blind. Blinding is important to maintaining prognostic balance as the study progresses, as it helps to minimize a variety of biases, such as placebo effects or co-interventions. Empirical evidence of bias exists in trials where blinding was not utilized or was ineffective.[10,11]

Five important groups should be blinded, when feasible: patients, clinicians, data collectors, outcome adjudicators, and data analysts [Table 1]. Frequently readers will see the terms “double-blind” or “triple-blind.” These terms may be confusing, and it is preferable to state exactly which groups are blinded in the course of a trial.[12] In surgical trials it is often impossible to blind the surgeon, but it may be feasible to blind patients, and is almost always feasible to blind data collectors and outcome assessors.

In the REDUCE trial,[6] several important groups were blinded. The control group received a placebo, so patients should have been blind to treatment allocation. Although not explicitly stated, it appears that clinicians were blinded, as the authors describe efforts to prevent unblinding by adjusting the PSA level of those patients on dutasteride to compensate for the expected reduction in this marker. The central pathology assessors were blinded to treatment allocation. It is unclear whether data collectors or analysts were blinded to treatment allocation. Overall, however, it does appear that the investigators made a reasonable attempt to blind important groups within the conduct of this trial.

| Table 1: Groups to consider blinding during conduct of randomized controlled trial and potential biases prevented[2] |
|-----------------|--------------------------------------------------|
| **Group**      | **Potential bias minimized**                      |
| Patients       | Placebo effect                                   |
| Clinicians     | Avoid co-intervention                            |
| Data collectors | Differential data collection                     |
| Outcome adjudicators | Differential assessment of whether an outcome of interest occurred |
| Data analysts  | Decisions that introduce bias during data analysis |

**Were the experimental and control groups prognostically balanced at the end of the study?**

**Was follow-up complete?**

In order to assure that both experimental and control groups are balanced at the end of a trial, complete follow-up information on each patient enrolled is important. Unfortunately, this is rarely the case at the close of a trial. Therefore, it is important to understand to what extent follow-up was incomplete. Incomplete follow-up can introduce bias into the trial results if loss to follow-up is not random. For example, consider a trial of medical versus surgical therapy for benign prostate hyperplasia. If patients randomized to medical therapy do poorly, and seek care outside of the trial (lost to follow-up), then the outcome of the study may be biased.

There is no specific rate of loss to follow-up, which is considered “acceptable” for a RCT. The loss to follow-up rate must be considered in context of the rate of outcome events for the trial. For example, consider a trial where the outcome is death from prostate cancer. If 30% of subjects die, and the loss to follow-up rate is 4%, then the loss to follow-up rate is much smaller than the outcome rate, and any influence of nonrandom loss to follow-up is likely to be small. However, what if only 5% of subjects in the trial died? In this case, a loss to follow-up rate of 4% is much more worrisome, as this is very close to the rate of outcome events and could dramatically change the trial results if loss to follow-up is not random.

In the REDUCE trial, the authors note that 1393 men never underwent prostate biopsy after enrolling in the trial.[6] In addition, it appears that 785 men did not undergo a second prostate biopsy as mandated by the study protocol. For the purposes of illustration, we will focus on the group of men who were enrolled in the trial but never underwent prostate biopsy. In the dutasteride group, 744 of the 4105 (18.1%) men randomized never had a prostate biopsy, whereas in the placebo group 649 of the 4126 (15.7%) men randomized never had a prostate biopsy (P=0.004). In terms of the outcome rate, 659 of the 4105 (16.1%) men randomized to dutasteride were diagnosed with prostate cancer, whereas 858/4126 (20.8%) of men randomized to placebo were diagnosed with prostate cancer. Notably, the event rate (prostate cancer diagnosis) and the lost to follow-up rate (no biopsy) are close, which may be concerning.

To further illustrate why the loss to follow-up rate is relatively high, consider the worst case scenario [Table 2]. Based on an intention to treat analysis, the relative risk of prostate cancer diagnosis over 4 years in the treatment group was 0.77 (23% relative risk reduction). In the worst case scenario, we assume that all subjects lost to follow-up on treatment experienced the outcome of interest (prostate cancer diagnosis), and none of the subjects lost to follow-up in the placebo arm experienced the outcome of interest.
Table 2: Potential impact of loss to follow-up and worst case scenario

|                   | Dutasteride | Placebo |
|-------------------|-------------|---------|
| No. randomized    | 4105        | 4126    |
| No. (%) lost to follow-up (no prostate biopsy) | 744 (18.1)  | 649 (15.7) |
| No. (%) diagnosed with prostate cancer      | 659 (16.1)  | 858 (20.8) |
| RR assuming no cancer in those lost to follow-up | 0.161/0.208 = 0.77 (23% relative risk reduction) |
| No. (%) with prostate cancer-worse case scenario | 1403 (34.2)  | 858 (20.8) |
| RR-worst case scenario | 0.342/0.208 = 1.64 (64% relative risk increase) |

This would give us 1403/4105 (34.2%) of men randomized with cancer in the dutasteride arm, and 858/4126 (20.8%) of men randomized with cancer in the placebo arm. In the worst case scenario, men on dutasteride would increase their risk of prostate cancer (RR 1.64). Obviously, it is highly improbable that all of the men randomized to dutasteride and subsequently lost to follow-up had prostate cancer, but this example illustrates the point that because the absolute reduction in risk of prostate cancer is small (5%), a relatively small imbalance in cancer diagnoses among those lost to follow-up could dramatically alter the estimated treatment effect of dutasteride on prostate cancer prevention.

Were patients analyzed in the groups to which they were randomized?

During a trial, not all patients receive the intervention to which they are assigned. Sometimes they are not adherent with therapy, or sometimes patients may even receive the experimental treatment instead of remaining in the control arm, a phenomenon known as crossover. When investigators remove these patients from analysis, the balance in prognosis created by randomization is undermined, and bias may be introduced into the estimate of the treatment effect. Use of the intention-to-treat principle helps to minimize bias created by nonadherence and/or crossover. The intention to treat principle holds that patients should be analyzed in the group to which they are randomized, regardless of whether they received the assigned intervention. Investigators frequently use the term intention-to-treat analysis, although this term can be misleading. Investigators often also present a per protocol analysis, which typically includes only those patients who actually received the intervention to which they were randomized and have complete follow-up.

In the REDUCE trial, the investigators do not report any crossover between groups. However, some patients in the trial were excluded from the analysis after randomization, which may conflict with the intention to treat principle. Among those excluded were 20 patients who did not receive any medication or placebo (nonadherence), the exclusion of which is likely not consistent with the intention to treat principle. An additional 53 men were excluded because they had a positive baseline biopsy on central review after randomization. If this central review was blinded to allocation, which is implied but not explicitly stated by the methods section, then this post-randomization exclusion may still be consistent with the intention-to-treat principle. Finally, 38 men had no baseline biopsy review and were excluded, which is likely not consistent with the intention-to-treat principle. The analysis of those men who underwent randomization and had at least one biopsy or were lost to follow-up (no biopsy), described as the efficacy population by the authors, is likely closest to the intention to treat principle. However, this does not minimize the potential bias introduced by loss to follow-up, as discussed previously. The “restricted crude rate” of prostate cancer diagnosis, which receives the most focus by the authors is the outcome among men who had at least one biopsy, and is probably best described as a per protocol analysis. The results of these two analyses (“restricted crude rate” and “crude rate”) are fairly similar, which is reassuring, but does not decrease the potential effect of significant loss to follow-up.

Was the trial stopped early?

At times, trials are stopped early and reported because of positive, large treatment effects. However, early termination may introduce bias secondary to chance deviations from the “true effect” of treatment which would decrease if the trial was continued to completion. Small trials and those with few outcome events are particularly prone to this bias if stopped early. For this reason, critical readers of the urology literature should interpret trials terminated early with caution. In the case of the REDUCE trial, it appears that the trial went to completion, so this is not a concern in terms of the validity of the trial.

Having completed our review of validity criteria, we must decide whether the results are likely to be valid before proceeding to answer our second question (What are the results?). As is common, validity in the case of the REDUCE trial is not an all-or-none phenomenon, but must be assessed on a continuum. Are the study methods strong and likely to yield an unbiased estimate, or are weaknesses present that will significantly bias the estimated treatment effect? It appears that the REDUCE trial was randomized and likely concealed, and known prognostic factors were balanced, suggesting participants started the trial with a similar prognosis. Prognostic balance was likely maintained during the study, with appropriate blinding of subjects, clinicians, and outcome assessors. At study’s completion, the question of prognostic balance is less certain because of a relatively high rate of loss to follow-up. Overall the study is well conducted and the methods are strong, but we must remember potential for bias introduced by loss to follow-up.
WHAT ARE THE RESULTS?

How large was the treatment effect?

Having decided that the study methods are reasonably strong, we continue reading to discover the estimate of the treatment effect. Note that any clinical trial only provides a point estimate of the treatment effect; it is not possible to know the true effect of any intervention. In this trial, the outcome is categorical: the presence or absence of prostate cancer. We can therefore use a $2 \times 2$ table [Table 3] to analyze the results of the study.

The authors report the main results of the study in a single sentence: "During the 4 years of the study, 659 of the 3305 men in the dutasteride group (19.9%) and 858 of the 3424 men in the placebo group (25.1%) received a diagnosis of prostate cancer, representing an absolute risk reduction with dutasteride of 5.1 percentage points."[6] From this sentence, we can construct a $2 \times 2$ table [Table 3] and calculate a variety of measures of the estimated treatment effect size.

The ARR is the absolute difference in event rates between the two groups: 25.1% - 19.9%, or 5.1% (due to rounding), as noted by the authors. Another way to express the results is the number needed to treat (NNT), which is simply the reciprocal of the ARR: 1/0.051, or 19.5, which we may round to 20. In other words, 20 men would need to be treated for 4 years with dutasteride to prevent one case of prostate cancer. Another way to express the results which is commonly used is the risk reduction, which is the experimental event rate divided by the control event rate, or in this case 0.199/0.251, or 0.793. The reduction in relative risk (RRR) is calculated as 1 - RR, or 1 - 0.793 = 20.7%. Empirical data suggest that when results are presented as RRR instead of ARR providers are more likely to recommend an intervention.[16,17] When we read the remainder of the results section of the REDUCE trial, the authors present various RRR of prostate cancer, which are higher than our calculations based on the main results in the first sentence (20.7% in our calculations vs 22.8% RRR for the “restricted crude rate”).[6] While it is not completely transparent how the authors calculate these RRRs, it is likely that some adjustment is made for removal of men from the cohort for prostate cancer diagnosis, and for additional biopsies outside of the trial protocol.[6]

How precise was the estimate of the treatment effect?

Recalling that the observed treatment effect is only an estimate of the true effect of the intervention, we would like to have some measure of the uncertainty surrounding the treatment estimate. This precision is usually communicated with a 95% confidence interval (CI). The width of the 95% CI depends on several factors, but the sample size is the most important. Trials with larger sample sizes will generally have smaller CIs than trials with fewer participants. In general, the 95% CI may be interpreted as the range of possible outcomes that will include the true effect size 95% of the time.[2] In the REDUCE report, the authors provide 95% CIs around all of the RRR, but we would also like to know the uncertainty around our ARR and NNT. Online calculators are readily available to assist with this task.[18] Using this calculator, we can determine that the 95% CI around the ARR (5.1%) is 3.1%-7.1%. Likewise, the calculator will provide us with 95% CIs for the NNT: 14-32. This suggests that as few as 14 men, or as many as 32 men, would require treatment for 4 years with dutasteride to prevent one case of prostate cancer.

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

Were the study patients similar to my patient?

Having accepted the trial results as likely valid, and decided that a benefit to treatment exists, we must then consider how the results of the trial can be applied to patient care. First, we must decide if our patient is similar enough to the study patients to apply the trial results. If our patient meets all of the inclusion and exclusion criteria for the study (ie, would have been eligible for the trial), then we believe the trial results will likely apply. In this case, our patient described in the scenario appears to meet eligibility criteria for the REDUCE trial, so we believe the results will be applicable. Not all patients will meet eligibility criteria for a given trial; if this is the case, we must decide if there is a compelling reason why the trial results would not apply to our patient.[2]

Were all patient-important outcomes considered?

The REDUCE trial considered a number of additional patient-centered outcomes, in addition to prostate cancer incidence. Several prostatic hyperplasia-related outcomes were evaluated, including prostate size, prostate symptoms, acute urinary retention, and prostate-related surgery. For most of these outcomes, treatment with dutasteride was superior to placebo. There was no difference in overall survival during the study period. One potentially patient-important outcome that is not reported is the cost of therapy. However, overall we can conclude that

---

Table 3: Prostate cancer diagnosis in men treated with 4 years of dutasteride or placebo and undergoing at least one biopsy[6]

| Outcome of interest (prostate cancer) | Present | Absent | Totals |
|--------------------------------------|---------|--------|--------|
| Treatment (dutasteride)              | 659 (a) | 2646 (b) | 3305   |
| Control (placebo)                    | 858 (c) | 2566 (d) | 3424   |
| Control event rate (CER) = c/(c+d)  | 858/3424 = 0.251 |
| Experimental event rate (EER) = a/(a+b) = 659/3305 = 0.199 |
| Absolute risk reduction (ARR) = CER-EER = 0.251-0.199 = 0.052 |
| Number needed to treat (NNT) = 1/ARR = 1/0.052 = 19 |
| Relative risk (RR) = EER/CER = 0.199/0.251 = 0.793 |
| Relative risk reduction (RRR) = 1-RR = 1-0.793 = 0.207 |
the authors did a good job reporting all major patient-
important outcomes.

Are the likely treatment benefits worth the harms and costs?

Finally, we have to weigh the potential treatment benefits with the potential harms and costs, while incorporating our patient’s values and preferences. Side effects related to loss of libido and erectile dysfunction were more common in the dutasteride group, and similar to that seen in previous trials of dutasteride for benign prostatic hyperplasia. There was an increase in a cardiac adverse event composite endpoint for dutasteride, although this effect had not been previously observed for dutasteride. Unlike an RCT of finasteride vs placebo for prostate cancer risk reduction, there was no overall increase in diagnosis of high-grade prostate cancer (which may have been an artifact in the finasteride trial). Interestingly, 11 more cases of Gleason 8-10 prostate cancer were detected in the treatment arm during years 3-4 of the REDUCE trial, which was statistically significant (P = 0.003). However, it is difficult to draw firm conclusions from a relatively small number of cases in this large trial. The lack of cost data makes it difficult to counsel our patient regarding the financial burden of 4 years of therapy with dutasteride.

CASE RESOLUTION

Your patient returns to clinic to discuss the findings of your reading. You describe the results of the REDUCE trial, as well as what you learned about the Prostate Cancer Prevention Trial of finasteride. At this time, the patient has minimal voiding symptoms, and has decided that he is not interested in starting medication solely for disease prevention, given the cost and potential side effects. If he begins to have urinary symptoms, he says that he would prefer to use dutasteride over an alpha blocker to treat his symptoms for the added potential benefit of prostate cancer prevention. You set up an appointment in 6 months to repeat a PSA and rectal examination.

CONCLUSION

Consumers of the urological literature should critically appraise articles regarding therapy. Three broad criteria should be assessed, including the validity of the results, the magnitude and precision of the treatment effect, and the applicability of results to patient care. Evidence-based clinical practice will lead toward higher quality patient care, and should be sought by all practicing urologists.

ACKNOWLEDGMENTS

The content of this review is based in part on concepts introduced in the Users’ Guides to the Medical Literature and the User’s Guide to the Urological Literature.

REFERENCES

1. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: What it is and what it isn’t. BMJ 1996;312:712.
2. Guyatt G, Rennie D, Meade MO, and 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.
3. 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.
4. 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;177:1090-5.
5. 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.
6. Andriole GL, Bostwick DG, Brawley OW, Gomella LG, Marberger M, Montorsi F, et al. Effect of dutasteride on the risk of prostate cancer. N Engl J Med 2010;362:1192-202.
7. Krupski TL, Dahm P, Fesperman SF, Schardt CM. How to Perform a Literature Search. J Urol 2008;179:1264-70.
8. Andriole G, Bostwick D, Brawley O, Gomella L, Marberger M, Tindall D, et al. Chemoprevention of prostate cancer in men at high risk: Rationale and design of the reduction by dutasteride of prostate cancer events (REDUCE) trial. J Urol 2004;172:1314-7.
9. Schulz SF, 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.
10. 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.
11. Lijmer JG, Mol BW, Heisterkamp S, Bonsel 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.
12. Devereaux PJ, Manns BJ, Ghali WA, Quan H, Lachet C, Montori VM, et al. Physician interpretations and textbook definitions of blinding terminology in randomized controlled trials. JAMA 2001;285:2000-3.
13. Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. BMJ 1999;319:670-4.
14. Ferguson DD, Aaron SD, Guyatt G, Hébert P. Post-randomisation exclusions: The intention to treat principle and excluding patients from analysis. BMJ 2002;325:652-4.
15. Montori VM, Devereaux PJ, Adhikari NK, Burns KE, Egger CH, Briel M, et al. Randomized trials stopped early for benefit: A systematic review. JAMA 2005;294:2203-9.
16. Forrow L, Taylor WC, Arnold RM. Absolutely relative: How research results are summarized can affect treatment decisions. Am J Med 1992;92:121-4.
17. Naylor CD, Chen E, Strauss B. Measured enthusiasm: Does the method of reporting trial results alter perceptions of therapeutic effectiveness? Ann Intern Med 1992;117:916-21.
18. Tandberg D. Calculator for the confidence interval around the difference of 4 years of therapy with dutasteride.
19. Thompson IM, Goodman PJ, Tangen CM, Lucia MS, Miller GJ, Ford LG, et al. The influence of finasteride on the development of prostate cancer. N Engl J Med 2003;349:215-24.
20. Pinsky P, Paresthes F, Ford L. Estimating rates of true high-grade disease in the prostate cancer prevention trial. Cancer Prev Res (Phila) 2008;1:182-6.
21. Redman MW, Tangen CM, Goodman PJ, Lucia MS, Colman CA Jr, Thompson IM. Finasteride does not increase the risk of high-grade prostate cancer: A bias-adjusted modeling approach. Cancer Prev Res (Phila) 2008;3:174-81.
22. Scales CD Jr, Preminger GM, Keitz SA, Dahm P. Evidence based clinical
practice: A primer for urologists. J Urol 2007;178:775-82.

23. Bajammal S, Dahm P, Scarpero HM, Orovan W, Bhandari M. How to use an article about therapy. J Urol 2008;180:1904-11.

How to cite this article: Stewart SB, Dahm P, Scales CD. How to appraise the effectiveness of treatment. Indian J Urol 2011;27:525-31.

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