



A practical approach to evidence-based dentistry: III

How to appraise and use an article about therapy

Romina Brignardello-Petersen, DDS, MSc;
Alonso Carrasco-Labra, DDS, MSc, PhD(c);
Michael Glick, DMD; Gordon H. Guyatt, MD, MSc;
Amir Azarpazhooh, DDS, MSc, PhD, FRCD(C)

THIRD IN A SERIES

In the 2 previous articles in this series, we introduced the process of evidence-based dentistry¹ and how to search for evidence to inform clinical practice.² In this article, we will explain how to use a research report to inform clinical decisions pertaining to questions of therapy. We will introduce and describe the basic concepts for understanding randomized controlled trials (RCTs), and we will explain how to critically appraise such studies. In subsequent

articles in this series, we will describe how to use other types of study designs.

CLINICAL QUESTIONS OF THERAPY

Dental practitioners spend most of their time administering treatments to their patients. A therapy or


ABSTRACT

Background and Overview. Dental practitioners spend most of their time administering treatments. To ensure that their clinical decisions are informed by the best available evidence, dental practitioners need to be skilled in critically appraising studies addressing therapy issues. Randomized controlled trials offer the optimal study design to inform decisions regarding therapy. The critical appraisal of randomized controlled trials involves assessing the risk of bias, results, and applicability. In this article, the authors present these concepts and provide guidance for this type of appraisal.

Practical Implications. Dentists who wish to inform their clinical decisions regarding therapy and prevention questions can use these guidelines to decide what type of studies to search, define the specific question of interest to search efficiently for these studies, and critically appraise an article about therapy or prevention.

Key Words. Evidence-based dentistry; therapy; prevention; randomized controlled trials; critical appraisal. JADA 2015;146(1):42-49

<http://dx.doi.org/10.1016/j.adaj.2014.11.010>

 Supplemental material is available online.

Dr. Brignardello-Petersen is a lecturer, Faculty of Dentistry, University of Chile, Santiago, Chile, and a doctoral student, Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada.

Dr. Carrasco-Labra is an instructor, Faculty of Dentistry, University of Chile, Santiago, Chile, and a doctoral student, Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada.

Dr. Glick is the dean, School of Dental Medicine, University at Buffalo, The State University of New York, Buffalo, NY. He also is the editor of the Journal of the American Dental Association.

Dr. Guyatt is a distinguished professor, Department of Clinical Epidemiology and Biostatistics, and a joint member, Department of Medicine, McMaster University, Hamilton, Ontario, Canada.

Dr. Azarpazhooh is an assistant professor, Dental Public Health and Endodontics, Faculty of Dentistry; and an assistant professor, Clinical Epidemiology and Health Care Research, Institute of Health Policy, Management and Evaluation, Faculty of Medicine, University of Toronto, Ontario, Canada. He also is the head of Endodontics, Mount Sinai Hospital, Toronto, Ontario, Canada. Address correspondence to Dr. Azarpazhooh at Biological and Diagnostic Sciences, 124 Edward St., Room 515C, Toronto, Ontario, Canada M5G 1G6, e-mail amir.azarpazhooh@dentistry.utoronto.ca.

This article has an accompanying online continuing education activity available at: <http://jada.ada.org/ce/home>.

Copyright © 2015 American Dental Association. All rights reserved.

BOX 1

Clinical scenario.

You referred a patient to an oral and maxillofacial surgeon for surgical extraction of an impacted mandibular third molar that has been causing the patient to have repeated episodes of pericoronitis. Your patient returned from the surgery consultation session and told you that the surgeon explained the risk of a postoperative infection and how to avoid it. The surgeon gave your patient the option to use either chlorhexidine gel or chlorhexidine mouth rinse. Your patient now is asking for your opinion regarding which of these 2 options may be more effective. You are not sure, so you decide to search for evidence from a clinical study to answer this question.

treatment can be defined as “any intervention, which may include prescribing drugs, performing surgery, or counseling, that is intended to improve the course of disease once it is established.”³ Many of the clinical questions that arise in clinical practice have to do with the effectiveness of treatments or interventions.

As described by the authors of a previously published article in this series,² therapy questions can be stated using the Population, Intervention, Comparison, Outcomes (PICO) framework.⁴ The population is the patients who are to receive the intervention, the intervention is the treatment of interest, the comparison is the reference to which we are comparing the intervention, and the outcomes are the health consequences that depend on the intervention. The comparison can be a different treatment or no treatment at all. Table 1 shows 2 examples of therapy questions and their corresponding questions in the PICO framework: 1 of them has no treatment as the comparison and the other has an alternative treatment.

WHAT STUDY DESIGN BEST ADDRESSES QUESTIONS OF THERAPY?

At the level of primary studies, RCTs represent the optimal study design to address questions of therapy. An RCT is an experiment assessing a medical treatment in patients.⁵ In an RCT, participants are allocated randomly into 2 or more groups that are treated equally except for the intervention the participants receive. After the intervention is applied, investigators follow patients over a prespecified time and measure outcomes, ideally those that are important to patients. If the study has been well designed and implemented, we can attribute differences that arise to the treatment under investigation (Figure).³

One RCT in dentistry, conducted by Hita-Iglesias and colleagues,⁶ addressed whether chlorhexidine gel or chlorhexidine rinse was more effective in preventing alveolar osteitis following third-molar extraction. After the surgery was performed, the investigators randomly allocated patients to receive either the gel or the rinse for 1 week, and evaluated patients on the third and

seventh days postextraction to determine the presence of alveolar osteitis.

RCTs are the best type of study design to determine the effectiveness of an intervention because they minimize bias—a systematic deviation from the underlying truth⁷—by ensuring that patients in the intervention and control groups are similar with respect to factors that determine whether the outcome of interest will occur. If well designed, RCTs also control events that occur after randomization and most aspects of the course of events, such as how the outcomes are measured.⁸ Therefore, clinicians should aim to inform their clinical decisions regarding therapy using individual RCTs or, even better, systematic reviews of RCTs. In this article, we will focus on how to use stand-alone RCTs, and in subsequent articles in this series, we will describe how to use systematic reviews.

BOX 2

The study you found.

During your search, you did not identify any summary or systematic review; however, you did find a randomized controlled trial (RCT) that seems to answer your question. The investigators of this RCT addressed whether postoperative chlorhexidine gel or chlorhexidine rinse was effective for preventing alveolar osteitis after extracting mandibular third molars.⁶ You read the abstract of this RCT, which indicated that the researchers recruited 73 patients and followed them for 1 week after the third-molar extraction, and it seemed as though chlorhexidine gel was more effective than the mouth rinse. However, you decide that you need to read the entire article to review more details before accepting its results and decide whether the results are applicable to your patient.

CRITICALLY APPRAISING AN RCT TO INFORM CLINICAL DECISIONS

The process of using an article from the dental literature to inform clinical decisions involves assessing the risk of bias, the results, and the applicability of the results.⁹ Below, we describe each of these 3 steps.

1. How serious is the risk of bias? The extent to which a study's results are likely to be correct for the sample of patients enrolled¹⁰ depends on how well the study was designed and conducted. Investigators of RCTs strive to ensure that determinants of the outcome of interest (factors such as age, sex, and disease severity, which we call prognostic factors) other than the treatment under investigation are similar between the groups being compared at the start of the study, and that these determinants remain similar throughout the study.⁸ Only if investigators

ABBREVIATION KEY. PICO: Population, Intervention, Comparison, Outcomes. RCT: Randomized controlled trial.

TABLE 1

Examples of therapy questions and the PICO* framework.				
CLINICAL QUESTION	POPULATION	INTERVENTION	COMPARISON	OUTCOMES
Should I prescribe chlorhexidine gel or chlorhexidine mouth rinse to a patient undergoing third-molar extractions to avoid postoperative infection?	Patients undergoing third-molar extractions	Chlorhexidine gel	Chlorhexidine mouth rinse	Postoperative infections (alveolar osteitis, surgical site infection)
Should I prescribe antibiotic prophylaxis for patients with diabetes undergoing endodontic treatment?	Patients with diabetes who will receive endodontic treatment	Antibiotic prophylaxis	No treatment (that is, no administration of antibiotic prophylaxis)	Postoperative infection, other postoperative complications
What analgesic should I prescribe to my patients to manage pain after dental extractions?	Patients undergoing dental extractions	Ibuprofen	Acetaminophen	Postoperative pain

* PICO: Population, Intervention, Comparison, Outcomes.

achieve and maintain prognostic balance can we be sure that any differences in outcomes are owing to the intervention (and not to bias introduced by the prognostic factors). Table 2¹¹⁻¹⁶ presents the aspects to consider when assessing the risk of bias of an RCT addressing a question of therapy.

1a. Did the intervention and control groups start with the same prognosis? One key aspect of a study that can help answer a question regarding therapy is whether the groups were balanced with regard to prognostic factors at the beginning of the study. Investigators can achieve this balance through their control of how the patients are allocated to the intervention and control groups.

Randomization assigns patients to the intervention or control groups by chance.⁸ The goal is to ensure that both known and unknown prognostic factors are distributed similarly in the intervention and control groups and, thus, avoid bias.^{9,17} This is why, when therapy questions are to be answered, the use of RCTs is superior to the use of observational studies: when the decision of what treatment to provide to a given patient is left to the dentist (as is done in observational studies), it is likely that patient characteristics may influence the choice of therapy. For instance, because of its ease of use, dentists may prefer to use chlorhexidine mouth rinse rather than gel in older patients who have a greater risk of experiencing subsequent infection, and to use the chlorhexidine gel in younger patients who have a lower risk of experiencing subsequent infection.

Even a well-prepared (typically, computer-generated) randomization schedule does not ensure random allocation. If those enrolling patients are aware of the treatment (intervention or control) to which the next patient will be allocated, they may make choices that undermine randomization.¹⁸ For example, if the next scheduled patient is an older person, and a member of the research staff who is responsible for allocation believes that the assigned treatment of gel is not optimal for the patient, the staff member may manipulate the

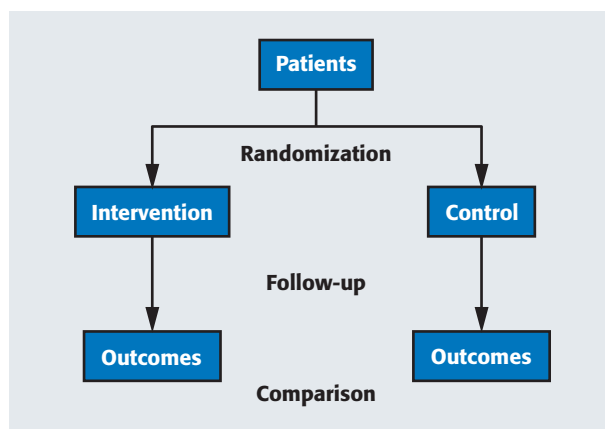


Figure. Randomized clinical trial.

allocation to ensure that the patient receives the mouth rinse.

To prevent this manipulation of the randomization schedule, those recruiting patients should not be informed about to which group the next patient will be allocated. This strategy is called “allocation concealment.”¹⁹ The authors of 1 study found that the investigators of trials in which allocation concealment was either inadequate or unclear reported treatment effects approximately 40% larger than those reported by studies with adequate allocation concealment.²⁰

Particularly if the sample size is small, even concealed randomization may fail to do its job of ensuring prognostic balance. It is important, therefore, to check whether the baseline characteristics of the patients in both groups are similar.⁹

In summary, concealed random allocation with evidence that patients in the intervention and control groups started with the same prognosis reassures us that we are likely to obtain unbiased estimates of treatment effect.

TABLE 2

Critically appraising the risk of bias in an article about therapy.*		
ASPECT	EXAMPLE	EXPLANATION
Did intervention and control groups start with the same prognosis?		
Were patients randomized?	"The sites presenting class II furcation lesions were randomly assigned, by a computer-generated list, to receive PDT [photodynamic therapy] or non-activated laser/only photo-sensitizer, both following [scaling and root planing]." ¹¹	Authors described that the sites were allocated randomly to the treatment groups and how the randomization sequence was generated. Examples of appropriate methods for generating the randomization sequence are random number tables, computer generators, and coin tossing. Inappropriate methods are those that do not produce true randomization, such as assigning patients to the groups on the basis of their date of birth or admission, or according to their record numbers. ¹²
Was randomization concealed?	"For each center, 16 consecutively numbered, opaque, sealed envelopes containing a note with the treatment (8 for each treatment) were made and placed in a larger envelope. For each patient, an independent person at each center randomly drew an envelope and handed it to Dentist B. This was repeated until 16 patients at each center were included." ¹³	Authors described using consecutively numbered, opaque, sealed envelopes, which were handed to the clinician by an independent person. This is an appropriate method for concealing the allocation. Other adequate methods include the use of sequentially numbered drug containers, and—by far the best of all—central allocation (telephone, web based, or pharmacy controlled). Allocation schedules or lists, envelopes without safeguards, or alternation are not appropriate for concealing allocation. ¹²
Were patients in the study groups similar with respect to known prognostic factors?	"There were no significant differences between groups for age, gender, duration of DM, glycemic status and category of DM regimen." ^{14†}	The authors presented a table with the baseline characteristics (potential prognostic factors) of the patients per group. When assessing this aspect, it is important to determine whether all relevant prognostic factors were considered.
Was prognostic balance maintained as the study progressed?		
To what extent was the study blinded?	"The entire study was blinded. The prophylaxis paste cups used had silver/blank lidstock and were only identified by a letter on the lidstock. The groups were not known by the examiners or patients. The examiner was in a different section of the building and the study coordinator gave the paste to the hygienist in yet another location of the building." ¹⁵	Authors described that their study was blinded. They also mentioned that the patients, the person administering the treatment (hygienist), and the examiners (outcome assessors) were blinded, and they described how this was achieved. Ideally, the authors should have mentioned that the data analyst was blinded as well. If blinding was not possible, it is necessary to judge the extent that this could have influenced the outcome measurement. ¹²
Were the groups prognostically balanced at the study's completion?		
Was follow-up complete?	"Analyses of the dropouts revealed no differences in TMJ pain, physical functioning, emotional functioning, or demographic data compared to the patients who completed the study." [‡] The information provided in the study flowchart shows that the number of patients lost to follow-up was 1 in one group (patients refused participation) and 3 in the other (1 patient refused participation and 2 moved from the area) at 10 weeks after the intervention. ¹³	Authors described the number and reasons for losing patients to follow-up. They also assessed whether these patients were different from those who continued in the study and found no differences. This showed that the risk of bias owing to incomplete follow-up was low. Other methods to assess this are to check whether the number of patients lost to follow-up and the reasons are similar between groups, whether the proportion of patients lost to follow-up is high enough to change the results if they were not missing, or to perform data imputation and draw conclusions on the basis of the results. ¹²
Were patients analyzed in the groups to which they were randomized?	"Trial outcomes were analysed by intention to treat. Per-treatment and per-protocol analyses of trial outcomes were also done for comparison." ¹⁶	Authors mentioned that they performed an intention-to-treat analysis. They described that they also did a per-protocol analysis to compare the results of both.
Was the trial stopped early?	"[I]t was determined that 16 subjects per group would be necessary to provide an 80% power with an alpha of 0.05. ... Thirty-eight subjects ... were selected from the population referred to the Periodontal Clinic of Guarulhos University." ¹⁴	Authors described the calculation of the target sample size, and later they mentioned the number of patients recruited, from which it can be inferred that they did not stop the trial early. The authors of the trials that have been stopped early owing to some data-dependent process usually mention this.
* Sources: Luchesi and colleagues, ¹¹ Higgins and colleagues, ¹² Christidis and colleagues, ¹³ Santos and colleagues, ¹⁴ Neuhaus and colleagues, ¹⁵ and Kelleher and colleagues. ¹⁶		
† DM: Diabetes mellitus.		
‡ TMJ: Temporomandibular joint.		

1b. Was prognostic balance maintained as the study progressed? Awareness on the part of dentists administering the intervention, patients who receive the care, or those measuring the outcomes in the study of whether patients are receiving intervention or control treatment can influence their behavior. For instance, the judgment of whether osteitis occurs involves some subjectivity, and it is possible that an outcome assessor who favors using gel over wash will be more likely to diagnose osteitis in those receiving the wash.²¹ Blinding (also known as masking) refers to investigators ensuring that patients, clinicians, and those collecting data and adjudicating whether an outcome has occurred are unaware of whether patients are receiving the intervention or the control.⁹ Typically, trials are described as double-blinded, which may mean that none of the patients, clinicians, data collectors, or outcome assessors are aware of what treatment the patient was receiving.²¹ The term “double-blind” does not leave us confident that all 4 groups are unaware of allocation, and optimal reporting will inform us whether this is the case.⁹

The method to achieve blinding in RCTs is the use of a placebo. A placebo is a treatment that resembles the intervention of interest but has no biological effect.⁸ Most placebos are pills and are administered in the same way as the real drug. Many of the treatments administered in dentistry, however, are procedures. In this case, although doing so is ethically arguable, investigators of the most rigorously designed trials could use a sham procedure to blind patients.²²

Many times, however, it is not possible to blind patients or clinicians. Fortunately, lack of blinding will not always result in bias. For instance, it may not be possible to blind the clinician because 2 materials with different appearance and techniques of use are being compared. If, however, patients and outcome assessors are blinded, and there are no other treatments that affect the outcome that the dentists may administer differentially to intervention and control patients (known as cointervention), not blinding the dentist may not affect the results. Therefore, when evaluating whether prognostic balance was maintained as the study progresses, not only do we have to assess whether the groups of interest were blinded, but also we must consider whether any lack of blinding may have caused any bias.²¹

1c. Were the groups prognostically balanced at the completion of the study? Strategies to maintain the prognostic balance as the study proceeds include following up on all patients, analyzing them in the group to which they were allocated, and completing the trial as planned.⁹

There are circumstances in which researchers are not able to follow up and measure the outcome for some

patients (referred to as lost to follow-up). Patients lost to follow-up may well have different prognoses from those who remain until the end of the study,⁹ and hence, the similarity of prognosis in intervention and control groups may be compromised. Ideally researchers should know the outcomes of all participants in a trial (or at least know the reasons why some patients were lost to follow-up).

If the proportion of patients lost to follow-up is so small that including their results in the data analysis would not change the overall results even if all of them had the best or worst outcome, and if the reasons for losing those patients are reported in the trial and are not potentially related to the outcome (for example, a patient moved to another city and could not attend the follow-up visits), then the risk of bias does not increase materially.

Sometimes patients do not receive the intervention as intended because they do not adhere to instructions. Patients may even receive the treatment meant for the patients allocated to the other group. If nonadherence to a treatment is related to a prognostic factor, excluding these patients from the analysis and using data only from patients who adhered to the treatment (that is, doing a per-protocol analysis) will destroy the prognostic balance achieved and maintained through all the previous strategies. To avoid this, researchers use the “intention-to-treat” principle; that is, they analyze the data from each patient in the group to which the patient was allocated. In this way, they get an unbiased estimate of the effect of the intervention at the level of adherence observed in the trial.²³

Randomization only can ensure prognostic balance when the sample size is large; when the sample size is small, chance can result in large prognostic imbalance. As a result, early in an RCT when sample sizes are small, results may not be indicators of the future overall results even if the RCT otherwise is designed meticulously.²⁴ Therefore, it can be misleading if investigators stop a trial early on the basis that they have seen a large effect.⁹ Both simulations and empirical data show that, on average, the results of trials that were stopped early because of a perceived benefit overestimate treatment effects, and the smaller the number of patients randomized as well as the number of outcome events observed at the time the trial is stopped, the larger the overestimate of effect.²⁵ Thus, a final criterion to determine whether the intervention and control groups were balanced prognostically at the end of the study is to assess whether the trial was stopped early.

In summary, prognostic balance is maintained up to the completion of the study if there are few participants who are lost to follow-up, if the authors performed an intention-to-treat analysis, and if the trial was completed as planned.

TABLE 3

Critically appraising the results of an article about therapy.		
EXAMPLE	HOW LARGE WAS THE TREATMENT EFFECT?	HOW PRECISE WAS THE ESTIMATE OF TREATMENT EFFECT?
"Compared to controls, the number of decayed teeth was significantly less in both the Xyl-2X* condition (relative risk [RR], 0.30; 95% confidence interval [CI] 0.13 to 0.66; P = 0.003) and the Xyl-3X† condition (RR, 0.50; 95% CI 0.26 to 0.96; P = 0.037)." ²⁶	The relative risk of the intervention groups compared with the control group was 0.3 and 0.5. This means that patients receiving the interventions had 0.3 times or 0.5 times the risk of caries than patients in the control group, which corresponds to a relative risk reduction of 70% and 50%, respectively. For both groups, this seems to be a large enough effect to consider administering this treatment in practice.	The confidence interval (CI) in 1 intervention group was 0.13 to 0.66. Because it is likely that both extremes of the CI would lead to a similar clinical action, this CI could be considered to be precise or narrow enough. On the other hand, the CI in the other intervention group was 0.26 to 0.96. The upper limit reflects only a 4% relative risk reduction in the risk of experiencing caries when using the intervention, which may not be an effect large enough so that a dentist would administer it. Because the extremes of the CI would lead to different clinical decisions, this CI is wide or not precise.
"There was, however, a statistical difference in mean attachment loss between the operated and unoperated (contralateral) canines (mean difference, 0.5 mm; 96% CI, 0.4-0.7; P < 0.001)." ²⁷	The difference in clinical attachment loss between the groups was 0.5 millimeters. A one-half mm of difference in attachment level seems to be very small, especially considering that the measurement error for this outcome has shown to be around 1 mm. ²⁸ Therefore, the magnitude of these results does not seem to be big enough to support the administration of the intervention.	The CI of the mean difference was 0.4 to 0.7 mm. Both extremes represented values of differences of attachment loss that had a small magnitude, therefore, this was a precise CI that provided the reader with more confidence in the small effect of the intervention, even accounting for the uncertainty of the results.
* Xyl-2X: 2 xylitol [4.00-g] doses and 1 sorbitol dose twice a day. † Xyl-3X: 3 xylitol [2.67-g] doses 3 times per day.		

2. **What are the results?** After assessing the magnitude of the risk of bias, clinicians must consider the results—in particular, the magnitude and the precision of the treatment effects—and the implications for their patient care. Table 3²⁶⁻²⁸ presents examples of the assessment of the results of an article about therapy.

2a. **How large was the treatment effect?** The researchers may analyze and present results of their RCTs in different ways, depending on the type of outcome of interest. When the outcome of interest is measured on a continuous scale (or numeric scale), such as probing depth and degree of trismus in millimeters, results usually are presented as differences in means between 2 groups. In this way, the mean of the outcome in the intervention group is subtracted from the mean of the outcome in the control group, and the numerical difference then would be attributed to the intervention. In this case, a value of 0 means that there is no difference between the intervention and control groups. Thus, the further the estimate of effect is from 0, the bigger the magnitude of the treatment effect.

On the other hand, the outcome of interest may be dichotomous (that is, the presence or absence or a condition or event), such as the presence of periodontal disease or trismus. With binary outcomes, investigators can present their results as the difference between the proportions in 2 groups. Imagine a trial comparing 2 drugs that prevent pain after dental extractions, and the outcome of interest is the presence or absence of pain. If 25% of the patients in the intervention group and 50% of the patients in the control group experience pain, the

BOX 3

Your assessment of the risk of bias of the randomized controlled trial you identified.

With respect to initial prognostic balance, you find that patients were randomized adequately and that the randomization schedule was likely to be concealed in an appropriate way. Nevertheless, there were some differences between groups in the distribution of men and women and in the distribution of smokers. With respect to maintaining prognostic balance as the study progressed, only the surgeons were blinded to the interventions the participants received. With respect to prognostic balance at the study completion, the authors did not describe how many, if any, patients were lost to follow-up in each treatment group. Your judgment leads you to consider that the lack of blinding of some participants and the lack of information regarding the losses to follow-up potentially may have biased the results, so you decide to keep reading this study with caution (see the supplemental Table⁶ for more details; available online at the end of this article).

risk difference is 25%. This also is referred to as absolute risk reduction. As for the mean difference, a value of 0 for the risk difference means that there is no difference between the intervention and control groups, and the further the estimate of effect is from 0, the bigger the magnitude of the treatment effect.

Another way to compare these proportions is to express 1 effect relative to the other—as the proportion of events in 1 group divided by the proportion of events in the other. In the example described previously, the relative effect (or relative risk) is $0.25 / 0.50 = 0.50$; patients receiving the intervention are at one-half of the risk of having pain as patients in the control group. From the relative risk we can derive the relative risk reduction, which reflects how much the risk of the

outcome increased or decreased in the intervention group relative to the control group. Because relative risk is best expressed as a percentage, in this example, the relative risk reduction would be 50%. When using relative measures of effect, a value of 1 means that the proportion of events is the same in both groups; the further the estimate of effect is from 1, the larger the treatment effect.

How large the effect is depends on the clinical context. For example, a difference of 2 mm in clinical attachment level gain when comparing an intervention with a control would be interpreted as a large magnitude by most clinicians. When considering another outcome, such as the difference in pain reduction using a visual analog scale of 100 mm, a difference of 2 mm between the groups would be interpreted as a small effect. Therefore, the magnitude of the effect reflects its importance to patients.²⁹

2b. How precise was the estimate of the treatment effect? Because of the influence of chance on results, researchers never can be completely sure of the true effect estimate. This is why they not only use point estimates such as the mean difference and the absolute and relative risks, but also complement them with estimates that express the degree of uncertainty in this point estimate. The estimate most commonly used to assess the precision of the results is the confidence interval (CI). A confidence interval (CI) is a plausible range of values within which the true value actually is likely to lie, given the data observed in a study.³⁰ A 95% CI range means that if the study was performed 100 times, the result would be within this range 95 times. The narrower the confidence interval, the more confident the researchers are of the estimate of effect. The interpretation of how wide or narrow a CI is depends—as it does in the case of the magnitude effect—on the clinical context.³⁰

BOX 4

Your assessment of the results of the randomized controlled trial you identified.

You find that patients who received chlorhexidine gel have 0.3 times the risk of developing alveolar osteitis than patients who received chlorhexidine mouth rinse. The 95% confidence interval of this relative risk ranged from 0.08 to 1.02. After looking at these numbers, you conclude that even though the magnitude of the treatment effect shows a large reduction in the risk of alveolar osteitis when using the gel, these results are not precise enough to claim that the gel is more effective than the mouth rinse (see the [supplemental Table⁶](#) for details; available online at the end of the article).

3. How can I apply the results to patient care?

Finally, clinicians should consider the extent to which the results of a study are applicable to their particular context. Factors clinicians should consider include whether the patients in the study have the same characteristics as the patients to whom they will apply the

results, how important the outcomes measured in the study are to their patients, whether the study investigators have measured all outcomes that are important to patients, and whether administering the intervention will result in more benefits than harms, costs, or both.

3a. Were the study patients similar to my patients?

In reading through a study, clinicians should evaluate the extent to which the patients studied are similar to the patients in their own practice and, hence, the extent to which results are applicable to their own patients (in other words, how generalizable are the results). When informing their clinical decisions with evidence from RCTs, clinicians should look at the selection criteria used for recruiting patients into a study and the description of the patients in the results. If the characteristics of the included patients are similar to the patients in their practice, clinicians could apply the results of the study in their practice. If there are characteristics that differ, such as age or severity of disease, clinicians should consider the likelihood that the effect of the intervention could be different if applied in their practice.⁹

3b. Were all patient-important outcomes considered? Ideally, the investigators of a study whose results can inform a clinical decision should have considered all patient-important outcomes. A patient-important outcome is any outcome that may alter a patient's inclination to choose a particular treatment, such as a reduction in symptoms, improvement in quality of life, treatment effect duration, or adverse effects.²⁹ Examples of patient-important outcomes in dentistry are tooth loss, pain, swelling, and the longevity of a restoration.

3c. Are the likely treatment benefits worth the potential harms and costs? Finally, when considering whether to administer a treatment to patients, clinicians should consider its benefits, harms, and costs.³¹ There is no treatment that has no undesirable consequence, even if the only downside is the time required for administration.³² Therefore, when considering using a treatment that seems to be beneficial, clinicians should put in the balance both the benefits and all potential downsides. The more inclined the balance is to 1 side, the easier it will be to make a decision.

BOX 5

Your assessment of the applicability of the randomized controlled trial you identified.

After assessing the applicability of the results, you conclude that even though the patients in the study were similar to your patient and there seem to be few harms associated with the use of either chlorhexidine gel or mouth rinse, the researchers did not consider all the infectious complications that could develop after the surgery. However, the authors did consider the most frequent infectious complication (see the [supplemental Table⁶](#) for details; available online at the end of the article). Thus, you decide that this evidence is still applicable to inform the choice your patient has to make.

CONCLUSION

Because of their design, RCTs are the best type of study to inform clinical decisions about therapy. However, clinicians should know how to appraise these studies to inform their decisions adequately. The critical appraisal of an RCT focuses on aspects of validity, results, and applicability. Clinicians should apply these guidelines to achieve the best possible results for their own practices. ■

BOX 6

What you say to your patient.

Despite the fact that the abstract of the study suggests that chlorhexidine gel may be more effective for preventing infectious complications after third-molar extractions than chlorhexidine mouth rinse, you inform your patient that there is no reason to believe that 1 of the options is better than the other, and you suggest that the patient should consider other aspects, such as the patient's preferences regarding the mode of administration and costs, when making a decision.

SUPPLEMENTAL DATA

Supplemental data related to this article can be found at <http://dx.doi.org/10.1016/j.ada.2014.11.010>.

Disclosure. None of the authors reported any disclosures.

- Brignardello-Petersen R, Carrasco-Labra A, Glick M, Guyatt GH, Azarpazhooh A. A practical approach to evidence-based dentistry: understanding and applying the principles of EBD. *JADA*. 2014;145(11):1105-1107.
- Brignardello-Petersen R, Carrasco-Labra A, Booth HA, et al. A practical approach to evidence-based dentistry: how to search for evidence to inform clinical decisions. *JADA*. 2014;145(12):1262-1267.
- Fletcher R, Fletcher S. Treatment. In: Fletcher R, Fletcher S, eds. *Clinical Epidemiology: The Essentials*. 4th ed. Baltimore, MD: Lippincott Williams & Wilkins; 2005:125-146.
- Guyatt GH, Meade MO, Richardson S, Jaeschke R. What is the question? In: Guyatt GH, Rennie D, Meade MO, Cook DJ, eds. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. 2nd ed. New York, NY: McGraw-Hill; 2008:17-28.
- Piantadosi S. Clinical trials as research. In: Piantadosi S. *Clinical Trials: A Methodologic Perspective*. 2nd ed. Hoboken, NJ: Wiley; 2005:9-28.
- Hita-Iglesias P, Torres-Lagares D, Flores-Ruiz R, et al. Effectiveness of chlorhexidine gel versus chlorhexidine rinse in reducing alveolar osteitis in mandibular third molar surgery. *J Oral Maxillofac Surg*. 2008;66(3):441-445.
- Glossary. In: Guyatt GH, Rennie D, Meade MO, Cook DJ, eds. *Users' Guide to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. 2nd ed. New York, NY: McGraw-Hill; 2008:769-808.
- Piantadosi S. Notation and terminology. In: Piantadosi S. *Clinical Trials: A Methodologic Perspective*. 2nd ed. Hoboken, NJ: Wiley; 2005:569-585.
- Guyatt GH, Straus S, Meade MO, et al. Therapy. In: Guyatt GH, Rennie D, Meade MO, Cook DJ, eds. *Users' Guide to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. 2nd ed. New York, NY: McGraw-Hill; 2008:67-86.
- Fletcher R, Fletcher S. Introduction. In: Fletcher R, Fletcher S, eds. *Clinical Epidemiology: The Essentials*. 4th ed. Baltimore, MD: Lippincott Williams & Wilkins; 2005:1-16.
- Luchesi VH, Pimentel SP, Kolbe MF, et al. Photodynamic therapy in the treatment of class II furcation: a randomized controlled clinical trial. *J Clin Periodontol*. 2013;40(8):781-788.

- Higgins J, Altman D, Sterne J. Assessing risk of bias in included studies. In: Higgins J, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions, Version 5.1.0 (updated March 2011)*. The Cochrane Collaboration; 2011. Available at: www.cochrane-handbook.org. Accessed November 19, 2014.

- Christidis N, Doepel M, Ekberg E, et al. Effectiveness of a pre-fabricated occlusal appliance in patients with temporomandibular joint pain: a randomized controlled multicenter study. *J Oral Facial Pain Headache*. 2014;28(2):128-137.
- Santos VR, Lima JA, Miranda TS, et al. Full-mouth disinfection as a therapeutic protocol for type-2 diabetic subjects with chronic periodontitis: twelve-month clinical outcomes: a randomized controlled clinical trial. *J Clin Periodontol*. 2013;40(2):155-162.
- Neuhaus KW, Millemann JL, Millemann KR, et al. Effectiveness of a calcium sodium phosphosilicate-containing prophylaxis paste in reducing dentine hypersensitivity immediately and 4 weeks after a single application: a double-blind randomized controlled trial. *J Clin Periodontol*. 2013;40(4):349-357.
- Kelleher J, Bhat R, Salas AA, et al. Oronasopharyngeal suction versus wiping of the mouth and nose at birth: a randomised equivalency trial. *Lancet*. 2013;382(9889):326-330.
- Roberts C, Torgerson D. Randomisation methods in controlled trials. *BMJ*. 1998;317(7168):1301.
- Torgerson DJ, Roberts C. Understanding controlled trials. Randomisation methods: concealment. *BMJ*. 1999;319(7206):375-376.
- Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. *Lancet*. 2002;359(9306):614-618.
- 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(5):408-412.
- Day SJ, Altman DG. Statistics notes: blinding in clinical trials and other studies. *BMJ*. 2000;321(7259):504-504.
- Piantadosi S. Contexts for clinical trials. In: Piantadosi S. *Clinical Trials: A Methodologic Perspective*. 2nd ed. Hoboken, NJ: Wiley; 2005:65-106.
- Montori V, Guyatt GH. Intention-to-treat principle. *Can Med Assoc J*. 2001;165(10):1339-1341.
- Weatley K, Clayton D. Be skeptical about unexpected large apparent treatment effects. *Control Clin Trials*. 2003;24(1):66-70.
- Bassler D, Briel M, Montori VM, et al. Stopping randomized trials early for benefit and estimation of treatment effects: systematic review and meta-regression analysis. *JAMA*. 2010;303(12):1180-1187.
- Milgrom P, Ly KA, Tut OK, et al. Xylitol pediatric topical oral syrup to prevent dental caries: a double-blind randomized clinical trial of efficacy. *Arch Pediatr Adolesc Med*. 2009;163(7):601-607.
- Parkin NA, Milner RS, Deery C, et al. Periodontal health of palatally displaced canines treated with open or closed surgical technique: a multicenter, randomized controlled trial. *Am J Orthod Dentofacial Orthop*. 2013;144(2):176-184.
- Espeland MA, Zappa UE, Hogan PE, Simona C, Graf H. Cross-sectional and longitudinal reliability for clinical measurement of attachment loss. *J Clin Periodontol*. 1991;18(2):126-133.
- Brignardello-Petersen R, Carrasco-Labra A, Sash P, Azarpazhooh A. A practitioner's guide to developing critical appraisal skills: what is the difference between clinical and statistical significance? *JADA*. 2013;144(7):780-786.
- Guyatt GH, Walter S, Cook D, Wyer P, Jaeschke R. Confidence intervals. In: Guyatt GH, Rennie D, Meade MO, Cook D, eds. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. 2nd ed. New York, NY: McGraw-Hill; 2008:99-107.
- Andrews JC, Schunemann HJ, Oxman AD, et al. GRADE guidelines, XV: going from evidence to recommendation—determinants of a recommendation's direction and strength. *J Clin Epidemiol*. 2013;66(7):726-735.
- Schunemann HJ. Guidelines 2.0: do no net harm—the future of practice guideline development in asthma and other diseases. *Curr Allergy Asthma Rep*. 2011;11(3):261-268.

eTABLE

Example of critically appraising an article about therapy.*	
Hita-Iglesias P, Torres-Lagares D, Flores-Ruiz R, et al. Effectiveness of chlorhexidine gel versus chlorhexidine rinse in reducing alveolar osteitis in mandibular third molar surgery. J Oral Maxillofac Surg 2008;66(3):441-445.	
1. How serious is the risk of bias?	
1a. Did the experimental and control groups begin the study with a similar prognosis?	
Were the patients randomized?	Yes. The study is described as a randomized clinical trial. The authors also described that they used a computer program to generate the allocation sequence.
Was randomization concealed?	Probably yes. The authors mentioned they used envelopes that indicated where the allocation assignment was recorded. These were opened right after the surgery. Ideally, they would have used opaque envelopes or the allocation sequence would have been kept in a different place such as a telephone central to prevent any misbehavior from researchers.
Were patients similar at baseline with respect to known prognostic factors?	Probably yes. The authors presented a comparison of the factors that they considered to be prognostically important in Table 1 of the article. These factors were fairly balanced between the gel and rinse group. The proportion of women and nonsmokers differed somewhat between the groups; however, it was thought that these differences were unlikely to bias the results.
1b. Was prognostic balance maintained as the study progressed?	
To what extent was the study blinded?	Surgeons were blinded because the allocation was done right after the surgery. Patients were not blinded because both antiseptics were administered in a different way and there was no mention of using placebos. There was no description of blinding of any of the other groups. Because the outcome had an explicit definition and the outcome assessor was trained, it was not very likely that the lack of blinding could have seriously biased the results.
1c. Were groups prognostically balanced at the study's completion?	
Was the follow-up complete?	Unclear. There was not enough information to make a judgment. The authors only reported the final number of participants in each group and did not mention whether this corresponded to all patients who were enrolled or whether some were lost to follow-up.
Were patients analyzed in the groups to which they were randomized?	Yes. The authors mentioned that they performed an intention-to-treat analysis. They also mentioned that there were no protocol violations as reported by patients, which would have made the results of both the intention-to-treat and the per-protocol analysis yield the same results.
Was the trial stopped early?	Probably not. There was no mention of stopping the trial early (which authors usually report). However, there were insufficient details to corroborate this, because there was no reporting of sample size calculation to assess whether the target sample size was reached.
2. What are the results?	
2a. How large was the treatment effect?	7.5% of the patients in the gel group and 25% of the patients in the rinse group developed alveolar osteitis. The relative risk was 0.3, indicating that patients who received chlorhexidine gel had 0.3 times the probability of developing alveolar osteitis than patients who received chlorhexidine rinse. The risk difference was 17.5%. Based on both measures of effect, the reduction in risk when using gel seemed to be of large magnitude.
2b. How precise was the estimate of the treatment effect?	Not precise. The authors did not include the confidence intervals, but with the numbers reported, it was possible to calculate the 95% confidence intervals for the relative risk. The confidence interval ranged between 0.08 and 1.02. Because this confidence interval suggested a very large benefit in 1 extreme, and a very small harm in the other, we considered the results to be imprecise (from a 92% reduction to a 2% increase in the incidence of alveolar osteitis).
3. How can I apply the results to my patient care?	
3a. Were the study patients similar to my patients?	Probably yes. The selection criteria described by the authors seemed to match the characteristics of most of the patients undergoing third-molar extraction who would be eligible to receive this intervention. However, depending on the particular context, the clinician should reassess this.
3b. Were all patient-important outcomes considered?	Probably not. The authors only measured alveolar osteitis and adverse effects. If they wanted to inform dental practice with the aim of preventing infectious complications like surgical site infection and other types of complications, then the study investigators did not provide information regarding all the relevant outcomes.
3c. Are the likely benefits worth the potential harms and costs?	Probably yes. The gel seemed to reduce the risk of developing alveolar osteitis, and it had similar tolerance and adverse effects (no adverse effects in both groups), thus the benefit-harm balances inclined toward the benefits. However, other potential issues such as burden of treatment and costs should be considered.
Conclusion: The results of the study are likely to be correct; however, even though the magnitude of the treatment effect showed a large reduction in the risk of developing alveolar osteitis when using the gel, these results were not precise enough to reach a sound conclusion. Aspects regarding applicability should be assessed further before making a decision regarding the administration of chlorhexidine gel over chlorhexidine rinse.	
* Source: Hita-Iglesias and colleagues. ⁶	