

Critical appraisal skills are essential to informed decision-making

Rahul Mhaskar^{1,3}, Patricia Emmanuel^{3,5}, Shobha Mishra⁴, Sangita Patel⁴,
Eknath Naik⁵, Ambuj Kumar^{1,2,3,5}

¹Center for Evidence-based Medicine and Health Outcomes Research, ²Department of Health Outcomes and Behavior, Moffitt Cancer Center, ³Clinical and Translational Science Institute, University of South Florida, ⁴Department of Preventive and Social Medicine, Government Medical College, Vadodara, India, ⁵Department of Internal Medicine, University of South Florida, College of Medicine, Tampa, Florida, USA.

Address for correspondence:

Dr. Ambuj Kumar, Division and Center for Evidence Based Medicine and Outcomes Research, Department of Internal Medicine, College of Medicine, University of South Florida, Clinical and Translational Science Institute, 12901 Bruce B Downs Blvd, MDC 3122, Tampa, FL 33612, E-mail: drambujkr@gmail.com

Abstract

Whenever a trial is conducted, there are three possible explanations for the results: a) findings are correct (truth), b) represents random variation (chance) or c) they are influenced by systematic error (bias). Random error is deviation from the 'truth' and happens due to play of chance (e.g. trials with small sample, etc.). Systematic distortion of the estimated intervention effect away from the 'truth' can also be caused by inadequacies in the design, conduct or analysis of a trial. Several studies have shown that bias can obscure up to 60% of the real effect of a healthcare intervention. A mounting body of empirical evidence shows that 'biased results from poorly designed and reported trials can mislead decision making in healthcare at all levels'. Poorly conducted and reported RCTs seriously compromise the integrity of the research process especially when biased results receive false credibility. Therefore, critical appraisal of the quality of clinical research is central to informed decision-making in healthcare. Critical appraisal is the process of carefully and systematically examining research evidence to judge its trustworthiness, value and relevance in a particular context. It allows clinicians to use research evidence reliably and efficiently. Critical appraisal is intended to enhance the healthcare professional's skill to determine whether the research evidence is true (free of bias) and relevant to their patients.

Key words: Critical appraisal, randomized controlled trials, decision-making

INTRODUCTION

A five-year-old male child presents with dry and scaly skin on his cheeks and was brought to the pediatric out-patient clinic by his parents. After initial evaluation, the young patient was diagnosed with atopic dermatitis. During the initial consultation, a junior resident mentions that a topical formulation of tacrolimus, an immunosuppressant currently marketed for the prevention of rejection after solid organ transplant, is a potential therapeutic agent for

atopic dermatitis. Nevertheless, the senior resident wants to know the evidence related to the safety and efficacy of topical tacrolimus in treating younger patients with atopic dermatitis. The junior resident enthusiastically performs an electronic search of the literature and finds a randomized controlled trial (RCT) conducted to determine the safety and efficacy of tacrolimus ointment in pediatric patients with moderate-to-severe atopic dermatitis.^[1] The junior resident also mentions that since this is an RCT, it should be considered reliable as it stands at a

How to cite this article:

Mhaskar R, Emmanuel P, Mishra S, Patel S, Naik E, Kumar A. Critical appraisal skills are essential to informed decision-making. *Indian J Sex Transm Dis* 2009;30:112-19.

DOI: 10.4103/0253-7184.62770

higher level in the hierarchy of the evidence pyramid [Figure 1]. However, the question now arises that because the trial claims to be a RCT, are the results from this study reliable, and whether the results are applicable to the young patient in question here?

THE NEED FOR CRITICAL APPRAISAL

Whenever a trial is conducted, there are three possible explanations for the results: a) findings are correct (truth), b) represents random variation (chance) or c) they are influenced by systematic error (bias). Random error is deviation from the 'truth' and happens due to play of chance (e.g. trials with small sample, etc.). Systematic distortion of the estimated intervention effect away from the 'truth' can also be caused by inadequacies in the design, conduct or analysis of a trial. Several studies have shown that bias can obscure up to 60% of the real effect of a healthcare intervention. A mounting body of empirical evidence shows that 'biased results from poorly designed and reported trials can mislead decision-making in healthcare at all levels'.^[2] Poorly conducted and reported RCTs seriously compromise the integrity of the research process especially when biased results receive false credibility. Therefore, critical appraisal of the quality of clinical research is central to informed decision-making in healthcare.

Critical appraisal is the process of systematically examining research evidence to judge its trustworthiness, its value and relevance in a particular context. It allows clinicians to use research evidence reliably and efficiently.^[3] Critical appraisal is intended to enhance the healthcare professional's skill to determine whether the research evidence is true (free of bias) and relevant to their patients. In this paper, we focus on the evaluation of an article (RCT) on a treatment intervention. The

same framework would be applicable to preventive interventions as well.

Three essential questions need to be asked when dealing with an article on therapeutic intervention.^[4]

- I. Are the results valid? Do the findings of this study represent the truth? That is, do the results provide an unbiased estimate of the treatment effect, or have they been clouded by bias leading to false conclusions?
- II. How precise are the results? If the results are unbiased, they need further examination in terms of precision. The precision would be better in larger studies compared with smaller studies.
- III. Are the results applicable to my patient? What are the patient populations, disease and treatments (including comparators) under investigation? What are the benefits and risks associated with the treatment? Do the benefits outweigh the harms?

- I. Are the results valid?
 - a. Did the study address a clearly focused issue?
Is there a clearly stated aim and research question? Ideally, a well-designed RCT should follow the acronym of PICOTS, which stands for patient, intervention, control, outcome, timing and setting, to formulate the research question. Is there a sound and understandable explanation of the population being studied (inclusion and exclusion criteria)? Is there an overview of which interventions are being compared? Are the outcomes being measured clearly stated with a rationale as to why these outcomes were selected for the study?
 - b. Was the assignment of participants to treatments randomized?

Randomization aims to balance the groups for known and unknown prognostic factors by allocating patients to two groups by chance alone. The aim is to minimize the probabilities of treatment differences attributed to chance and to maximize the attribution to treatment effects. Therefore, it is important that the intervention and control groups are similar in all aspects apart from receiving the treatment being tested. Otherwise, we cannot be sure that any difference in outcome at the end is not due to pre-existing disparity. If one group has a significantly different average age or social class make-up, this might be an explanation of why that group did better or worse. The best method to

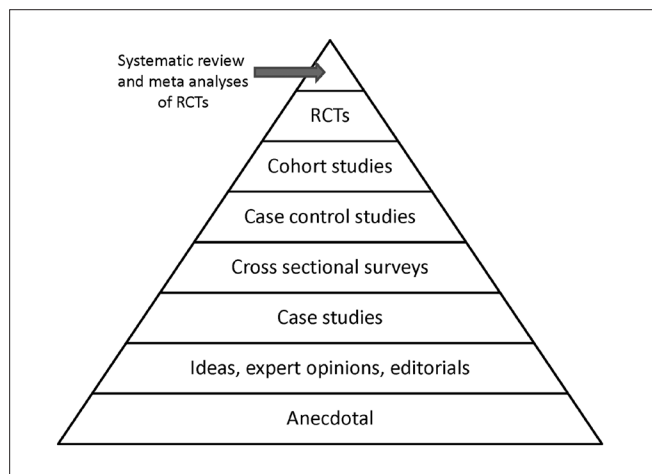


Figure 1: Evidence pyramid showing the hierarchy of evidence.^[13]

create two groups that are similar in all important respects is by deciding entirely by chance, into which group a patient will be assigned.^[5] This is achieved by random assignment of patients to the groups. In true randomization, all patients have the same chance as each other of being placed into any of the groups.^[4,5] Allocation concealment ensures that those assessing eligibility and assigning subjects to groups do not have knowledge of the allocation sequence. In order to ensure adequate allocation concealment, centralized computer randomization is ideal and is often used in multicenter RCTs. The successful randomization process will result in similar groups. It is important to note that randomization will not always produce groups balanced for known prognostic factors. A large sample size will potentially reduce the likelihood of placing individuals with better prognoses in one group.^[4,6]

- c. Were all the participants who entered the study properly accounted for at its conclusion? Description of dropouts and withdrawals: How do dropouts threaten validity?

Participants in a RCT are considered lost to follow-up when their status or outcomes are not known at the end of the study. Dropouts might happen due to several natural reasons (e.g. participant relocation). It is assumed that such dropouts will not necessarily be a substantial number. However, when the dropout rate is large (e.g. >10%), it is a reason for concern. Often the reason that patients are lost to follow up relates to a systematic difference in their prognosis compared with patients who continue with a study until the end (e.g., patients lost to follow-up may have more adverse events or worse outcomes than those who continue). Therefore, the loss of many participants may threaten internal validity of the trial. Furthermore, if loss to follow-up and drop outs is different between two study groups, it may result in missing data that can disrupt the balance in groups created by randomization.

The abovementioned questions should be adequate for screening purposes of the manuscript to help make a decision whether to continue assessing the article further. That is, if the answers to the first three questions are negative, it is not worth evaluating the findings, as the results from

this study would not be considered reliable to be used in decision-making.

II. How precise are the results?

- a. What are the results?

The readers of the paper should be attentive as to whether authors of the study adequately describe the data collection that was employed in the study. These should be clearly described and justified. All outcome measures should be referenced and their validity reviewed. If data was self-reported by patients, it would need to be verified in some way for maximum credibility.

Were the methods of analysis appropriate, clearly described and justified? Analysis should relate to the original aims and research questions. Choice of statistical analyses should be explained with a clear rationale. Any unconventional tests should be justified with references to validation studies.

An important issue to consider in the analysis of a RCT is whether the analysis was performed according to the intention to treat (ITT) principle. According to this principle, all patients are analyzed in the same treatment arm regardless of whether they received the treatment or not. Intention to treat (ITT) analysis is important because it preserves the benefit of randomization. If certain patients in either arm dropped out, were non-compliant or had adverse events are not considered in the analysis, it is similar to allowing patients to select their treatment and the purpose of randomization is rendered futile.

Alternatively, a 'per-protocol' or 'on treatment' analysis includes only patients who have sufficiently complied with the trial's protocol. Compliance covers exposure to treatment, availability of measurements and absence of major protocol violations and is necessary to measure the treatment-related harm (TRH) outcomes. Adhering to ITT analysis for assessment of TRH leads to biased estimates.^[7,8] In summary, analysis of benefits associated with treatments should be performed using the ITT principle and associated risks according to per protocol.

- b. What are the key findings?

The findings should answer the primary research question(s). Each outcome measure should be analyzed and its results presented with comparisons between the groups. Other issues of analysis include the magnitude

of the effect, which can be measured by relative and absolute risk differences, odds ratios and numbers needed to treat. The significance of any differences between the groups should be discussed, with *p*-values given to indicate statistical significance (<0.05 being the common threshold for significance). The confidence intervals should be presented to demonstrate the degree of precision of the results. Was the sub-group analyses preplanned or derived post hoc (derived from a fishing expedition after data collection)? For example, in the RCT mentioned in the clinical scenario, tacrolimus in children aged 6 to 17 years suffering from atopic dermatitis had a statistically significant difference compared with vehicle group for the outcome of Physician's Global Evaluation of clinical response.^[1] It is important to note that the study was adequately powered to address the primary question only. It would be incorrect if the investigators decided to address the issue of efficacy according to age groups (e.g. subgroup of 6 to 8 years old versus 9 to 17) or any other subgroups if it was not decided a priori. That is, subgroup analysis could be justified for hypothesis generation but not hypothesis testing. The validity of subgroup-treatment effects can only be tested by reproducing the results in future trials. Rarely are trials powered to detect subgroup effects. There is often a high false negative rate for tests of subgroup-treatment effect interaction when a true

interaction exists.^[9]

The authors should include the descriptive analysis of the data (mean, median, standard deviation, frequencies etc.) and not just the results of statistical tests used. Most often, results are presented as dichotomous outcomes (yes or no). Consider a study in which 25% (0.25) of the control group died and 20% (0.20) of the treatment group died. The results can be expressed in many ways as shown in Table 1.^[10]

- c. Are the results applicable to my patient? How will the results help me work with my patients/clients? Can the results be applied to the local population of my practice and clients? How similar is the study sample to your own clients? Are there any key differences that you would need to consider for your own practice? In the case of our five-year-old with atopic dermatitis, even if the trial was performed flawlessly with positive results, if the population studied was not similar then the evidence is not applicable. An example of this would be if the study was performed in patients with severe atopic disease who had failed all other regimens but our patient was naïve to any therapy. If the population was a different age range or ethnicity, it could also impact relevance. Other important considerations are whether you have the necessary skills or resources to deliver the intervention. Were all the important outcomes considered? Has the research covered the most important

Table 1: Essential terminologies and their interpretations

What is the measure?	What does it mean?
Relative Risk (RR) = observed risk of the outcome in the treatment group / observed risk of the outcome in the control group. In our example, the RR = 0.20/0.25 = 0.80	Relative risk is a ratio of the probability of the event occurring in the treatment group versus a non-treatment (placebo) group. An RR of 1 indicates: no difference between the two groups. An RR < 1 indicates: treatment reduces the risk of the outcome. An RR > 1 suggests that the treatment increased the risk of the outcome.
Absolute Risk Reduction (ARR) = risk of the outcome in the control group - risk of the outcome in the treatment group. In our example, the ARR = 0.25 - 0.20 = 0.05 or 5%	ARR indicates the decrease in risk of a given outcome in patients with the treatment in relation to risk of that outcome in individuals not receiving the treatment. An ARR of 0 suggests that there is no difference between the two groups indicating that treatment had no effect.
Relative Risk Reduction (RRR) = absolute risk reduction / risk of the outcome in the control group. In our example, the RRR = 0.05/0.25 = 0.20 or 20%	RRR indicates the reduction in the rate of the outcome in the treatment group relative to that in the control group.
Number Needed to Treat (NNT) = 1 / ARR. In our example, the NNT = 1 / 0.05 = 20	NNT is the number of patients needs to be treated in order to prevent one additional bad outcome.

Adopted from critical appraisal sheet University of Oxford Center for Evidence-based Medicine, 2005.^[10]

outcomes for your patients? If key outcomes were overlooked, do you need further evidence before changing your practice? Are the benefits of the intervention worth the harms and the costs? If the study does not answer this question, you will need to use your own judgment, taking into account your patients, all the stakeholders, yourself and your colleagues.

Absolute Risk Reduction (ARR) = risk of the outcome in the control group – risk of the outcome in the treatment group.

In our example, the ARR = 0.25 – 0.20 = 0.05 or 5%. Absolute Risk Reduction (ARR) indicates the decrease in risk of a given outcome in patients with the treatment in relation to risk of that outcome in individuals not receiving the treatment.

An ARR of '0' suggests that there is no difference between the two groups indicating that treatment had no effect.

Relative Risk Reduction (RRR) = absolute risk reduction / risk of the outcome in the control group.

In our example, the RRR = 0.05/0.25 = 0.20 or 20%. Relative Risk Reduction (RRR) indicates the reduction in the rate of the outcome in the treatment group relative to that in the control group.

Number Needed to Treat (NNT) = 1 / ARR.

In our example, the NNT = 1/ 0.05 = 20. Number Needed to Treat (NNT) is the number of patients needs to be treated in order to prevent one additional bad outcome.

In conclusion, the reporting of RCT can be plagued with numerous quality control issues. Consolidated Standards of Reporting Trials group (CONSORT) has developed various initiatives to improve the issues arising from inadequate reporting of RCTs. The main products of CONSORT are the CONSORT statement^[11] and CONSORT harms statement,^[12] which are an evidence-based, minimum set of recommendations for reporting RCTs. These offer a standard way for authors to prepare reports of trial findings, facilitating their complete and transparent reporting and aiding their critical interpretation.^[11] In essence there is a need to assess the quality of evidence; and, if adequate, establish the range of true treatment effect. Then, consider whether results are generalizable to the patient at hand, and whether the measured outcomes are relevant and important.

Finally, carefully review the patient's risk of TRH and related treatment benefit – risk ratio.^[6] We believe that methodologically assessing the strength of evidence and using it to guide treatment of each patient will certainly improve health outcomes.

Additional material

A critical appraisal worksheet (with permission from <http://www.cebm.net/index.aspx?o=1157>) is provided in the appendix section of the manuscript. We encourage the readers to assess the manuscript mentioned in the clinical scenario^[1] and critically appraise it using the worksheet (see appendix).

ACKNOWLEDGEMENT

This paper was supported in part by the Fogarty International Center/USNIH: Grant # 1D43TW006793-01A2-AITRP.

REFERENCES

1. Boguniewicz M, Fiedler VC, Raimer S, Lawrence ID, Leung DY, Hanifin JM. A randomized, vehicle-controlled trial of tacrolimus ointment for treatment of atopic dermatitis in children. *J Allergy Clin Immunol* 1998;102:637-44.
2. Moher D, Schulz KF, Altman D; CONSORT Group. The CONSORT Statement: Revised recommendations for improving the quality of reports of parallel-group randomized trials 2001. *Explore (NY)* 2005;1:40-5.
3. Burls A. What is critical appraisal? 2009.
4. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA* 1993;270:2598-601.
5. Jüni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. *BMJ* 2001;323:42-6.
6. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients? Evidence-Based Medicine Working Group. *JAMA* 1994;271:59-63.
7. Armitage. Exclusions, losses to follow-up and withdrawals in clinical trials. *Clinical trials*, Shapiro SH. New York: Marcel Dekker; 1983.
8. Gail MH. Eligibility exclusions, losses to follow-up, removal of randomized patients, and uncounted events in cancer clinical trials. *Cancer Treat Rep* 1985;69:1107-13.
9. Rothwell PM. Treating individuals 2. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. *Lancet* 2005;365:176-86.
10. Oxford, T.U.o. Critical appraisal sheet for RCT; 2009.
11. Egger M, Altman DG, Schulz KG. The CONSORT Statement: Explanation and Elaboration. *Ann Intern Med* 2002;136:926-7.
12. Ioannidis JP, Evans SJ, Gøtzsche PC, O'Neill RT, Altman DG, Schulz K, *et al.* Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004;141:781-8.
13. Sackett DL. Evidence-based medicine. How to practice and teach EBM. New York: Churchill Livingstone Inc; 2000. P. 136.

Source of Support: Fogarty International Center/
USNIH: Grant #1D43TW006793-01A2-AITRP.
Conflict of Interest: None declared.

Appendix

Therapy study: Are the results of the trial valid? (Internal Validity)

What question did the study ask?

Patients –

Intervention -

Comparison -

Outcome(s) -

1a. R- Was the assignment of patients to treatments randomised?	
What is best?	Where do I find the information?
Centralised computer randomisation is ideal and often used in multi-centred trials. Smaller trials may use an independent person (e.g, the hospital pharmacy) to “police” the randomization.	The Methods should tell you how patients were allocated to groups and whether or not randomisation was concealed.
This paper: Yes <input type="checkbox"/> No <input type="checkbox"/> Unclear <input type="checkbox"/> Comment:	
1b. R- Were the groups <u>similar</u> at the start of the trial?	
What is best?	Where do I find the information?
If the randomisation process worked (that is, achieved comparable groups) the groups should be similar. The more similar the groups the better it is. There should be some indication of whether differences between groups are statistically significant (ie. p values).	The Results should have a table of "Baseline Characteristics" comparing the randomized groups on a number of variables that could affect the outcome (ie. age, risk factors etc). If not, there may be a description of group similarity in the first paragraphs of the Results section.
This paper: Yes <input type="checkbox"/> No <input type="checkbox"/> Unclear <input type="checkbox"/> Comment:	
2a. A – Aside from the allocated treatment, were groups treated equally?	
What is best?	Where do I find the information?
Apart from the intervention the patients in the different groups should be treated the same, eg., additional treatments or tests.	Look in the Methods section for the follow-up schedule, and permitted additional treatments, etc and in Results for actual use.
This paper: Yes <input type="checkbox"/> No <input type="checkbox"/> Unclear <input type="checkbox"/> Comment:	
2b. A – Were all patients who entered the trial accounted for? – and were they analysed in the groups to which they were randomised?	
What is best?	Where do I find the information?
Losses to follow-up should be minimal – preferably less than 20%. However, if few patients have the outcome of interest, then even small losses to follow-up can bias the results. Patients should also be analysed in the groups to which they were randomised – ‘intention-to-treat analysis’.	The Results section should say how many patients were randomised (eg., Baseline Characteristics table) and how many patients were actually included in the analysis. You will need to read the results section to clarify the number and reason for losses to follow-up.
This paper: Yes <input type="checkbox"/> No <input type="checkbox"/> Unclear <input type="checkbox"/> Comment:	
3. M - Were measures <u>objective</u> or were the patients and clinicians kept “ <u>blind</u> ” to which treatment was being received?	
What is best?	Where do I find the information?
It is ideal if the study is ‘double-blinded’ – that is, both patients and investigators are unaware of treatment allocation. If the outcome is objective (eg., death) then blinding is less critical. If the outcome is subjective (eg., symptoms or function) then blinding of the outcome assessor is critical.	First, look in the Methods section to see if there is some mention of masking of treatments, eg., placebos with the same appearance or sham therapy. Second, the Methods section should describe how the outcome was assessed and whether the assessor/s were aware of the patients' treatment.
This paper: Yes <input type="checkbox"/> No <input type="checkbox"/> Unclear <input type="checkbox"/> Comment:	

What were the results?	
1. How large was the treatment effect?	
<p>Most often results are presented as dichotomous outcomes (yes or not outcomes that happen or don't happen) and can include such outcomes as cancer recurrence, myocardial infarction and death. Consider a study in which 15% (0.15) of the control group died and 10% (0.10) of the treatment group died after 2 years of treatment. The results can be expressed in many ways as shown below.</p>	
What is the measure?	What does it mean?
<p>Relative Risk (RR) = risk of the outcome in the treatment group / risk of the outcome in the control group.</p> <p>In our example, the RR = 0.10/0.15 = 0.67</p>	<p>The relative risk tells us how many times more likely it is that an event will occur in the treatment group relative to the control group. An RR of 1 means that there is no difference between the two groups thus, the treatment had no effect. An RR < 1 means that the treatment decreases the risk of the outcome. An RR > 1 means that the treatment increased the risk of the outcome.</p> <p>Since the RR < 1, the treatment decreases the risk of death.</p>
<p>Absolute Risk Reduction (ARR) = risk of the outcome in the control group - risk of the outcome in the treatment group. This is also known as the absolute risk difference.</p> <p>In our example, the ARR = 0.15 - 0.10 = 0.05 or 5%</p>	<p>The absolute risk reduction tells us the absolute difference in the rates of events between the two groups and gives an indication of the baseline risk and treatment effect. An ARR of 0 means that there is no difference between the two groups thus, the treatment had no effect.</p> <p>The absolute benefit of treatment is a 5% reduction in the death rate.</p>
<p>Relative Risk Reduction (RRR) = absolute risk reduction / risk of the outcome in the control group. An alternative way to calculate the RRR is to subtract the RR from 1 (eg. RRR = 1 - RR)</p> <p>In our example, the RRR = 0.05/0.15 = 0.33 or 33%</p> <p>Or RRR = 1 - 0.67 = 0.33 or 33%</p>	<p>The relative risk reduction is the complement of the RR and is probably the most commonly reported measure of treatment effects. It tells us the reduction in the rate of the outcome in the treatment group relative to that in the control group.</p> <p>The treatment reduced the risk of death by 33% relative to that occurring in the control group.</p>
<p>Number Needed to Treat (NNT) = inverse of the ARR and is calculated as 1 / ARR.</p> <p>In our example, the NNT = 1/ 0.05 = 20</p>	<p>The number needed to treat represents the number of patients we need to treat with the experimental therapy in order to prevent 1 bad outcome and incorporates the duration of treatment. Clinical significance can be determined to some extent by looking at the NNTs, but also by weighing the NNTs against any harms or adverse effects (NNHs) of therapy.</p> <p>We would need to treat 20 people for 2 years in order to prevent 1 death.</p>
2. How precise was the estimate of the treatment effect?	
<p>The true risk of the outcome in the population is not known and the best we can do is estimate the true risk based on the sample of patients in the trial. This estimate is called the point estimate. We can gauge how close this estimate is to the true value by looking at the confidence intervals (CI) for each estimate. If the confidence interval is fairly narrow then we can be confident that our point estimate is a precise reflection of the population value. The confidence interval also provides us with information about the statistical significance of the result. If the value corresponding to no effect falls outside the 95% confidence interval then the result is statistically significant at the 0.05 level. If the confidence interval includes the value corresponding to no effect then the results are not statistically significant.</p>	
Will the results help me in caring for my patient? (ExternalValidity/Applicability)	
<p>The questions that you should ask before you decide to apply the results of the study to your patient are:</p> <ul style="list-style-type: none"> • Is my patient so different to those in the study that the results cannot apply? • Is the treatment feasible in my setting? • Will the potential benefits of treatment outweigh the potential harms of treatment for my patient? 	

Evidence Based Medicine

- Level I: Evidence obtained from at least one properly designed randomized controlled trial.
- Level II-1: Evidence obtained from well-designed controlled trials without randomization.
- Level II-2: Evidence obtained from well-designed cohort or case-control analytic studies, preferably from more than one center or research group.
- Level II-3: Evidence obtained from multiple time series with or without the intervention. Dramatic results in uncontrolled trials might also be regarded as this type of evidence.
- Level III: Opinions of respected authorities, based on clinical experience, descriptive studies, or reports of expert committees.
-
- Level A : Consistent randomized controlled clinical trial, cohort study, all or none, clinical decision rule validated in different populations
- Level B : Consistent Retrospective cohort, exploratory cohort, ecological study, outcomes research, case control study or extrapolations from level A study
- Level C : Case series study or extrapolations from level B studies
- Level D : Expert opinion without explicit critical appraisal, or based on physiology, bench research or first principles
-
- Level A: Good scientific evidence suggests that the benefits of the clinical service substantially outweighs the potential risks. Clinicians should discuss the service with eligible patients.
- Level B: At least fair scientific evidence suggests that the benefits of the clinical service outweighs the potential risks. Clinicians should discuss the service with eligible patients.
- Level C: At least fair scientific evidence suggests that there are benefits provided by the clinical service, but the balance between benefits and risks are too close for making general recommendations. Clinicians need not offer it unless there are individual considerations.
- Level D: At least fair scientific evidence suggests that the risks of the clinical service outweighs potential benefits. Clinicians should not routinely offer the service to asymptomatic patients.
- Level I: Scientific evidence is lacking, of poor quality, or conflicting, such that the risk versus benefit balance cannot be assessed. Clinicians should help patients understand the uncertainty surrounding the clinical service.