

Critical Appraisal of Article on Therapy

What question did the study ask?

Population/problem: _____

Intervention: _____

Comparison: _____

Outcome: _____

Guide	Comment
Are the results Valid	
1. Was the assignment of patients to treatments randomized? And was the randomization list concealed?	
2. Was follow-up of patients sufficiently long and complete?	
3. Were patients analyzed in the groups which they were randomized?	
4. Were patients and clinicians kept "blind" to treatment?	
5. Were the groups treated equally, apart from the experimental treatment?	
6. Were the groups similar at the start of the trial?	
Are the valid results of this randomized study important?	
1. What is the magnitude of the treatment effect?	
2. How precise is this estimate of the treatment effect?	
Are these valid, important results applicable to our patient?	
1. Is our patient so different from those in the study that its results cannot apply?	
2. Is the treatment feasible in our setting?	
3. What are our patient's potential benefits and harms from the therapy?	
4. What are our patient's values and expectations for both the outcome we are trying to prevent and the treatment we are offering?	

How to Review an Article on Therapy

What question did the study ask?

Population/problem: _____

Intervention: _____

Comparison: _____

Outcome: _____

Guide	Comment
Are the results Valid	
1. Was the assignment of patients to treatments randomized? 2. And was the randomization list concealed?	<ul style="list-style-type: none"> • Randomization is important in balancing prognostic factors (both known & unknown) between treatment & control groups. • This methodology is the gold standard for clinical trials. • It is important that physicians who enter patients into a Randomized Clinical Trials (RCT) cannot influence which group their patient enters (control or experimental). The randomization should be concealed in some fashion – sealed envelopes or calling to remote site for assignment are two examples.
2. Was follow-up of patients sufficiently long and complete?	<ul style="list-style-type: none"> • Was the follow-up of patients sufficiently long to see a clinically important effect? Example – several weeks is fine for streptococcal pharyngitis, while years may be appropriate for chronic diseases, ie cancer, cardiac disease • Was there an acceptable loss of follow-up of patients? Greater than 20% is usually considered unacceptable.
3. Were patients analyzed in the groups which they were randomized?	<ul style="list-style-type: none"> • To preserve the value of randomization “intention to treat analysis” should be performed. • The subject is analyzed in the group which they were randomized.
4. Were patients and clinicians kept “blind” to treatment?	<ul style="list-style-type: none"> • Blinding both the clinician and patient to the treatment (or lack thereof) is ideal. • Sometimes patients and clinicians can’t be blinded, such as in surgical trials. • The most important blinding is that of the assessment of the outcomes of the study.
5. Were the groups treated equally, apart from the experimental treatment?	<ul style="list-style-type: none"> • The experimental and control group should be treated equally apart from the experimental treatment. They should have the same testing, the same number of follow-up visits, the same educational interventions other than experimental treatment.

<p>6. Were the groups similar at the start of the trial?</p>	<ul style="list-style-type: none"> • While randomization should make the groups similar, they may not be exactly equal. • Groups should be similar in all prognostically important ways. • If they are not similar, there should be some adjustment for potentially important prognostic factors carried out in the analysis phase. This can include stratification or multiple regression analysis.
<p>Are the valid results of this randomized study important?</p>	
<p>1. What is the magnitude of the treatment effect?</p>	<p>CER – control event rate EER – experimental event rate Relative risk reduction (RRR) = $(CER - EER)/CER$ Absolute risk reduction (ARR) = $CER - EER$ Number Needed to Treat (NNT) = $1/ARR$ *Note – there are formulas to calculate confidence intervals for each of the above measures. They are not included as you are not expected to be able to calculate them.</p>
<p>2. How precise is this estimate of the treatment effect?</p>	<ul style="list-style-type: none"> • Look for 95% confidence intervals. Confidence intervals are the measure of precision • The wider the confidence intervals, the less precise the measurement. This is relative.
<p>Are these valid, important results applicable to our patient?</p>	
<p>1. Is our patient so different from those in the study that its results cannot apply?</p>	<ul style="list-style-type: none"> • Our patient does not have to fit all the inclusion criteria of this study. • Consider whether our patient's sociodemographic features or pathobiology are so different from those in the study that its results are useless to us and our patient.
<p>2. Is the treatment feasible in our setting?</p>	<ul style="list-style-type: none"> • Is the treatment economically feasible and available in our geographic region?
<p>3. What are our patient's potential benefits and harms from the therapy?</p>	<ul style="list-style-type: none"> • There is always constant weighing of the treatment's potential benefits and harms.
<p>4. What are our patient's values and expectations for both the outcome we are trying to prevent and the treatment we are offering?</p>	<ul style="list-style-type: none"> • We must elicit our patient's preferences for both the outcome we are trying to prevent and the treatment we are offering.

HOW TO REVIEW AN ARTICLE ON DIAGNOSIS

What question did the study ask?

Population/problem: _____

Intervention: _____

Comparison: _____

Outcome: _____

Guides	Comments			
Are the results of this diagnostic study valid?				
1. Was there an independent “blind” comparison with a reference (gold) standard of diagnosis?				
2. Was the diagnostic test evaluated in an appropriate spectrum of patients (like those in whom it would be used in practice)?				
3. Was the reference standard applied regardless of the diagnostic test result?				
4. Was the test (or cluster of tests) validated in a second, independent group of patients?				
Are the valid results of this diagnostic study important?				
2x2 Table				
Diagnostic Test Result	Target Disorder	Total		
	Present	Absent		
	Positive	a	b	a + b
	Negative	c	d	c + d
	Total	a + c	b + d	a + b + c + d
1. Sensitivity = $a/(a+c)$ 2. Specificity = $d/(b+d)$ 3. Sensitivity = $a/(a+c)$				

<p>4. Likelihood Ratio for a positive test result = $LR+ = \text{sens}/(1-\text{spec})$</p> <p>5. Likelihood Ratio for a negative test result = $LR- = (1-\text{sens})/\text{spec}$</p> <p>6. Positive Predictive Value = $a/(a+b)$</p> <p>7. Negative Predictive Value = $d/(c+d)$</p> <p>8. Pre-test Probability (prevalence) = $(a+c)/(a+b+c+d)$</p> <p>9. Pre-test-odds = prevalence/(1-prevalence)</p> <p>10. Post-test odds = Pre-test odds x Likelihood Ratio</p> <p>11. Post-test Probability = Post-test odds/(Post-test odds + 1) + d</p> <p>12. Likelihood Ratio for a positive test result = $LR+ = \text{sens}/(1-\text{spec})$</p> <p>13. Likelihood Ratio for a negative test result = $LR- = (1-\text{sens})/\text{spec}$</p> <p>14. Positive Predictive Value = $a/(a+b)$</p> <p>15. Negative Predictive Value = $d/(c+d)$</p> <p>16. Pre-test Probability (prevalence) = $(a+c)/(a+b+c+d)$</p> <p>17. Pre-test-odds = prevalence/(1-prevalence)</p> <p>18. Post-test odds = Pre-test odds x Likelihood Ratio</p> <p>19. Post-test Probability = Post-test odds/(Post-test odds + 1)</p>	
<p>Can we apply this valid, important evidence about a diagnostic test in caring for our patient?</p>	
<p>1. Is the diagnostic test available, affordable, accurate, and precise in our setting?</p>	
<p>2. Can we generate a clinically sensible estimate of our patient's pre-test probability (from personal experience, prevalence statistics, or primary studies)?</p> <ul style="list-style-type: none"> • Are the study patients similar to our own? • Is it unlikely that the disease possibilities or probabilities have changed since this evidence was gathered? 	
<p>3. Will the resulting post-test probabilities affect our management and help our patient?</p> <ul style="list-style-type: none"> • Could it move us across a test- treatment threshold? • Would our patient be a willing partner in carrying it out? 	

HOW TO REVIEW AN ARTICLE ON DIAGNOSIS

What question did the study ask?

Population/problem: _____

Intervention: _____

Comparison: _____

Outcome: _____

Guides	Comments
Are the results of this diagnostic study valid?	
1. Was there an independent “blind” comparison with a reference (gold) standard of diagnosis?	<ul style="list-style-type: none"> • The gold standard is used to define the true state of the patient • The gold standard must be reasonable and doable, but does not have to be the perfect test. • Example – coronary angiography is considered the gold standard for the identification of coronary artery disease. One could argue that autopsy is the gold standard, but difficult to perform in live patients
2. Was the diagnostic test evaluated in an appropriate spectrum of patients (like those in whom it would be used in practice)?	<ul style="list-style-type: none"> • You want to see that patients with a wide spectrum of the target diagnosis (mild, moderate & severe) and patients with commonly confused diagnoses (e.g., diagnostic value of rales – include patients with congestive heart failure, chronic lung disease, pneumonia) are included. • Spectrum bias refers to the bias introduced by having too narrow a spectrum of patients (severity or other commonly confused diagnoses)
3. Was the reference standard applied regardless of the diagnostic test result?	<ul style="list-style-type: none"> • Every patient in the study should get both the reference standard and the diagnostic test being studied. • Beware of studies that only compare the reference standard and diagnostic test in patients who are positive for a screening test (can be either reference standard or

	diagnostic test. For example, a study examining the value of d-dimer in the diagnosis of deep vein thrombosis, only patients who had a positive venous ultrasound (reference standard) got a d-dimer.
4. Was the test (or cluster of tests) validated in a second, independent group of patients?	<ul style="list-style-type: none"> Ideally (especially for cluster of tests or clinical prediction rules), the diagnostic test should be validated in a second independent sample. This does not often occur.
Are the valid results of this diagnostic study important?	
See worksheet on test characteristics for formulas.	In addition for likelihood ratios – SpPin – LR+ > 10 SnNout – LR- < 0.1
Can we apply this valid, important evidence about a diagnostic test in caring for our patient?	
1. Is the diagnostic test available, affordable, accurate, and precise in our setting?	<ul style="list-style-type: none"> In order to be useful, a test needs to be affordable and available. It doesn't do your patient any good if they cannot afford the test or have to go to the North Pole to get it.
2. Can we generate a clinically sensible estimate of our patient's pre-test probability (from personal experience, prevalence statistics, or primary studies)? <ul style="list-style-type: none"> Are the study patients similar to our own? Is it unlikely that the disease possibilities or probabilities have changed since this evidence was gathered? 	
3. Will the resulting post-test probabilities affect our management and help our patient? <ul style="list-style-type: none"> Could it move us across a test- treatment threshold? Would our patient be a willing partner in carrying it out? 	

HOW TO REVIEW AN SYSTEMATIC REVIEW

What question did the study ask?

Population/problem: _____

Intervention: _____

Comparison: _____

Outcome: _____

Guide	Comments
Deciding whether the results are valid	
1. Is this a systematic review of randomized trials?	
2. Does it include a methods section that describes: (a) Finding and including all the relevant trials? (b) Assessing their individual validity	
3. Were the results consistent from study to study?	
4. Were individual patient data used in the analysis or aggregate data? (may important in meta-analysis)	
Are the valid results of the systematic review important?	
1. What is the magnitude of the treatment effect?	
2. How precise is the treatment effect?	
Will the results help me in caring for my patients?	
1. Is our patient so different from those in the study that its results cannot apply?	
2. Is the treatment feasible in our setting?	
3. What are our patient's potential benefits and harms form the therapy?	
4. What are our patient's values and expectations for both the outcome we are trying to prevent and the treatment we are offering?	

HOW TO REVIEW A SYSTEMATIC REVIEW

What question did the study ask?

Population/problem: _____

Intervention: _____

Comparison: _____

Outcome: _____

Guide	Comments
Deciding whether the results are valid	
1. Is this a systematic review of randomized trials?	<ul style="list-style-type: none"> • Systematic reviews (SRs) may go by other names including overview and meta-analysis. A meta-analysis is really a subset of SR where the results of many studies are statistically combined. • A SRs is like any other research – it has a focused question. Its data are the results of other studies instead of 'generating new data'. • While most SRs will only include randomized trials, some will include high quality studies, such as cohort studies. If the SR includes studies other than randomized trials, they should be analyzed separate. In other words a meta-analysis should not combine data from both randomized and non-randomized studies.
2. Does it include a methods section that describes: (a) finding and including all the relevant trials? (b) assessing their individual validity?	<ul style="list-style-type: none"> • A good SR will examine many sources for evidence including standard bibliographic databases, hand-searching journals, conference proceedings, theses, databanks of pharmaceutical firms and contacting 'experts' or authors of published articles. Really good SRs will also search the European literature and/or foreign language literature. • Each study that meets the inclusion criteria (or doesn't fail exclusion criteria) should be independently assessed by at least two reviewers. If there is a disagreement between reviewers, there should be some mechanism to resolve the differences. There are published criteria for reviewing and numerically rating studies.
3. Were the results consistent from study to study?	<ul style="list-style-type: none"> • While different studies may show differing quantitative effects, the studies should be consistent (homogeneous) with respect to overall benefit or harm. Beware of the SR that has some studies that show benefit, some no benefit and some clear-cut harm. These results would be very heterogeneous.

	<ul style="list-style-type: none"> • If the results are heterogeneous, the author should explain why such as differing study populations (by race, gender, and age), doses of medication, or duration of therapy.
4. Were individual patient data used in the analysis or aggregate data? (may important in meta-analysis)	<ul style="list-style-type: none"> • Combining individual patient data is preferable. It allows you to examine subgroups. Combining summary data does not offer the same flexibility. To use individual patient data you would need to go back to the original source of the data and obtain permission to use it.
Are the valid results of the systematic review important?	
1. What is the magnitude of the treatment effect?	<ul style="list-style-type: none"> • Many SRs are beginning to report their results with NNT. However, the reporting of Odds Ratio (OR) or Relative Risk (RR) is still very common. Both OR and RRR can be converted into NNT.
2. How precise is the treatment effect?	<ul style="list-style-type: none"> • Look for 95% confidence intervals. Confidence intervals are the measure of precision • The wider the confidence intervals, the less precise the measurement. This is relative
Will the results help me in caring for my patients?	
1. Is our patient so different from those in the study that its results cannot apply?	<ul style="list-style-type: none"> • Our patient does not have to fit all the inclusion criteria of this study. • Consider whether our patient's sociodemographic features or pathobiology are so different from those in the study that its results are useless to us and our patient.
2. Is the treatment feasible in our setting?	<ul style="list-style-type: none"> • Is the treatment economically feasible and available in our geographic region?
3. What are our patient's potential benefits and harms from the therapy?	<ul style="list-style-type: none"> • There is always constant weighing of the treatment's potential benefits and harms.
4. What are our patient's values and expectations for both the outcome we are trying to prevent and the treatment we are offering?	<ul style="list-style-type: none"> • We must elicit our patient's preferences for both the outcome we are trying to prevent and the treatment we are offering.