

Additional Evidence Appraisal Worksheets[†]

DIAGNOSTIC TEST— EVIDENCE APPRAISAL WORKSHEET

Citation:

Are the results of this diagnostic study valid?

Appraisal Criterion	Comments
<p>Was there an independent, masked comparison between the diagnostic test of interest and a “gold (reference) standard” diagnostic test?</p> <ul style="list-style-type: none"> • If not, describe what was done, the limitations of this approach, and the potential consequences for the study’s results. 	
<p>Was the diagnostic test evaluated in subjects with the range of presentation (i.e., different levels or stages) of the condition?</p> <ul style="list-style-type: none"> • If not, briefly describe the sample and discuss the potential consequence that this limited sample has for the study results. 	

[†]Adapted with permission from the Oxford Center for Evidence-Based Medicine (www.cebm.net).

Appraisal Criterion	Comments
<p>Did the investigators perform the “gold standard” diagnostic test on every subject regardless of the result from the diagnostic test of interest?</p> <ul style="list-style-type: none"> • If not, describe what was done, as well as the limitations of this approach. 	
<p>Was the test (or cluster of tests) evaluated in a second, independent group of subjects?</p> <ul style="list-style-type: none"> • If no, describe the limitations resulting from the lack of a comparison group. 	

Are the valid results of this diagnostic study important?

Appraisal Criterion	Comments
<p>What were the statistical findings of this study?</p> <ul style="list-style-type: none"> • When appropriate, use the calculation forms below to determine these values. 	<ul style="list-style-type: none"> • Tests of Association? P-values? Confidence Intervals? • Sensitivity? • Specificity? • Positive Predictive Value? • Negative Predictive Value? • Positive Likelihood Ratios? Confidence Intervals? • Negative Likelihood Ratio? Confidence Intervals?
<p>What is the meaning (application) of these statistical findings for your patient/client case?</p>	

Can you apply this valid, important evidence about a diagnostic test in caring for your patient/client?

Appraisal Criterion	Comments
Does the test sound appropriate for use (available, affordable, reliable, and valid) in your clinical setting?	
Are the study patients/clients similar to your own? <ul style="list-style-type: none"> • If not, how are they different? • Can you use this test in spite of these differences? 	
Can you generate a clinically sensible estimate of your patient’s/client’s pretest probability of the disorder (from personal experience, prevalence statistics, practice databases, or primary studies)? <ul style="list-style-type: none"> • Is it possible that the disease probabilities have changed since the data you are using were gathered? If so, how will you adjust your pretest probability? 	
Would the test and its results, including the posttest probabilities, help your patient/client? <ul style="list-style-type: none"> • If so, how? • If not, could the test and its results cause harm to the patient/client? 	
Does the test fit within your patient’s/client’s stated values or expectations? <ul style="list-style-type: none"> • If not, what will you do now? 	

Additional notes:

SAMPLE CALCULATIONS

From Holtby and Razmjou ²⁰		Target Disorder (Biceps Pathology & SLAP Lesion)		Totals
		Present	Absent	Totals
Diagnostic Test Result (Yergason's)	Positive (Pain in Bicipital Groove or GH Joint)	a 9	b 6	a+b 15
	Negative (No Pain in Bicipital Groove or GH Joint)	c 12	d 22	c+d 34
<i>Totals</i>		a+c 21	b+d 28	a+b+c+d 49

Sensitivity = $a/(a+c) = 9/21 = 43\%$

Specificity = $d/(b+d) = 22/28 = 79\%$

Likelihood Ratio for a Positive Test Result = $LR+ = Sens/(1-Spec) = 43\%/21\% = 2.05$

Likelihood Ratio for a Negative Test Result = $LR- = (1-Sens)/Spec = 57\%/79\% = 0.72$

Positive Predictive Value = $a/(a+b) = 9/15 = 60\%$

Negative Predictive Value = $d/(c+d) = 22/34 = 65\%$

Pretest Probability (prevalence) = $(a+c)/(a+b+c+d) = 21/49 = 43\%$

Pretest Odds = $Prevalence/(1-Prevalence) = 43\%/57\% = 0.75$

Posttest Odds = $Pretest Odds \times LR = 0.75 \times 2.05 = 1.54$

Posttest Probability = $Posttest Odds/(Posttest Odds + 1) = 1.54/2.54 = 61\%$

Holtby R, Razmjou H. Accuracy of the Speed's and Yergason's tests in detecting biceps pathology and SLAP lesions: Comparison with arthroscopic findings. *Arthroscopy*. 2004; 20(3): 231-236.

YOUR CALCULATIONS

		Target Disorder		Totals
		Present	Absent	Totals
Diagnostic Test Result	Positive	a	b	a+b
	Negative	c	d	c+d
<i>Totals</i>		a+c	b+d	a+b+c+d

PROGNOSIS—EVIDENCE APPRAISAL WORKSHEET

Citation:

Are the results of this prognosis study valid?

Appraisal Criterion	Comments
Was a defined, representative sample of subjects included in the study? • If not, what are the sample’s limitations and what are the potential consequences for this study’s results?	
Were the subjects assembled at a common (usually early) point in the course of their disorder? • If not, what are the implications of multiple starting points for this study’s results?	
Was the subject follow-up time sufficiently long to answer the question(s) posed by the research? • If not, what are the potential consequences of the follow-up time for the study’s results?	
Did all subjects originally enrolled complete the study? • If not, how many subjects were lost? • What, if anything, did the authors do about this attrition? • What are the implications of the attrition and the way it was handled with respect to the study’s findings?	
Were objective outcome criteria applied to the subjects in a masked (“blind”) fashion? • If not, what are the potential consequences for the study’s results?	

Appraisal Criterion	Comments
<p>If subgroups with different prognoses are identified, was there adjustment for important prognostic (risk) factors?</p> <ul style="list-style-type: none"> • If not, what should the investigators have included? • What are the potential consequences of the absence of adjustment for this study results? 	
<p>Was there validation in an independent group (“test set”) of patients?</p> <ul style="list-style-type: none"> • If not, what are the potential consequences for this study’s results? 	

Are the valid results of this prognosis study important?

Appraisal Criterion	Comments
<p>How likely are the outcomes over time?</p>	
<p>What do the results for the predictors or risk factors indicate (e.g., odds ratios, relative risk)?</p>	
<p>How precise are the prognostic estimates?</p>	<p>CI (%)</p> <p>*Use the table at the end of this worksheet to calculate confidence intervals if not provided</p>

Can you apply this valid, important evidence about prognosis in caring for your patient?

Appraisal Criterion	Comments
Were the study subjects similar to your patient/client? • If not, what are the differences and how do they affect your use of the study?	
Will this evidence make a clinically important impact on your conclusions about what to offer or tell your patient/client? • If yes, how so? • If not, why not?	

Additional notes:

***If you want to calculate a confidence interval around the event rate for one group:**

Clinical Measure	Standard Error (SE)	Typical Calculation of CI
Proportion (i.e., the rate of some prognostic event, etc.) where: n = the number of patients p = the proportion of these patients who experience the event	$\sqrt{\{p \times (1 - p)/n\}}$	If $p = 24/60 = 0.4$ (or 40%) and $n=60$ $SE = \sqrt{\{0.4 \times (1 - 0.4)/60\}}$ $= 0.063$ (or 6.3%) 95% CI is 40% \pm $1.96 \times 6.3\%$ or 27.6%–52.4%
n from your evidence: _____ p from your evidence: _____	$\sqrt{\{p \times (1 - p)/n\}}$	Your calculation: SE: _____ 95% CI:

INTERVENTION—EVIDENCE APPRAISAL WORKSHEET

Citation:

Are the results of this therapeutic trial valid?

Appraisal Criterion	Comments
<p>Did the investigators randomly assign subjects to treatment groups?</p> <ul style="list-style-type: none"> • If no, describe what was done. • What are the potential consequences of this assignment process for the study's results? 	
<p>Did the investigators know who was being assigned to which group prior to the allocation?</p> <ul style="list-style-type: none"> • If yes, what are the potential consequences of this knowledge for the study's results? 	
<p>Were the groups similar at the start of the trial?</p> <ul style="list-style-type: none"> • If not, what differences existed? • How might the differences between groups affect the results of this study? 	
<p>Did subjects know to which treatment group they were assigned?</p> <ul style="list-style-type: none"> • If yes, what are the potential consequences of the subjects' knowledge for this study's results? 	
<p>Did investigators know to which treatment group subjects were assigned?</p> <ul style="list-style-type: none"> • If yes, what are the potential consequences of the investigators' knowledge for this study's results? 	

<p>Were the groups managed equally, apart from the experimental treatment?</p> <ul style="list-style-type: none"> • If not, what are the potential consequences of this difference in management for the study's results? 	
<p>Was the subject follow-up time sufficiently long to answer the question(s) posed by the research?</p> <ul style="list-style-type: none"> • If not, what are the potential consequences of the follow-up time for the study's results? 	
<p>Did all subjects originally enrolled complete the study?</p> <ul style="list-style-type: none"> • If not, how many subjects were lost? • What, if anything, did the authors do about this attrition? • What are the implications of the attrition and the way it was handled with respect to the study's findings? 	
<p>Were all patients analyzed in the groups to which they were randomized (i.e., was there an intention-to-treat analysis)?</p> <ul style="list-style-type: none"> • If not, what did the authors do with the data from these subjects? • If the data were excluded, what are the potential consequences for this study's results? 	

Are the valid results of this randomized trial important?

Appraisal Criterion	Comments
<p>What were the statistical findings of this study?</p> <ul style="list-style-type: none"> • When appropriate, use the calculation forms below to determine these values. 	<ul style="list-style-type: none"> • Tests of Differences? P-values? Confidence Intervals? • Effect Sizes? P-values? Confidence Intervals? • Absolute and Relative Benefit Increases? P-values? Confidence Intervals?

Appraisal Criterion	Comments
	<ul style="list-style-type: none"> • Absolute and Relative Risk Reductions? P-values? Confidence Intervals? • Number Needed to Treat (Harm)? Confidence Intervals? • Other?
<p>What is the meaning (application) of these statistical findings for your patient/client’s case?</p>	
<p>Do these findings exceed a minimal clinically important difference?</p> <ul style="list-style-type: none"> • If not, will you still use this evidence? 	

Can you apply this valid, important evidence about an intervention in caring for your patient/client?

Appraisal Criterion	Comments
<p>Does the intervention sound appropriate for use (available, affordable) in your clinical setting?</p>	
<p>Are the study subjects similar to your patient/client?</p> <ul style="list-style-type: none"> • If not, how different? Can you use this intervention in spite of these differences? 	
<p>Do the potential benefits outweigh the potential risks of using this intervention with your patient/client?</p>	
<p>Does the intervention fit within your patient/client’s stated values or expectations?</p> <ul style="list-style-type: none"> • If not, what will you do now? 	

Additional notes:

SAMPLE CALCULATIONS—BENEFIT INCREASE

	+ Outcome	- Outcome
+ Intervention	A 72	B 55
- Intervention	C 40	D 78

CER = Control Group Event Rate = $c / (c + d) = 0.25 = 34\%$

EER = Experimental Group Event Rate = $a / (a + b) = 0.57 = 57\%$

		Relative Benefit Increase (RBI)	Absolute Benefit Increase (ABI)	Number Needed to Treat (NNT)
EER	CER	$\frac{EER - CER}{CER}$	EER - CER	1/ABI
57%	34%	56%	23%	4
		*95% CI	10.9-35.1%	3-9

***95% Confidence Interval (CI) on an NNT =**
1/(Limits on the CI of its ABI) =

$$\pm 1.96 \sqrt{\left[\frac{CER \times (1 - CER)}{\# \text{ Control Pts}} \right] + \left[\frac{EER \times (1 - EER)}{\# \text{ Exper Pts}} \right]} =$$

$$\pm 1.96 \sqrt{\left[\frac{0.34 \times 0.66}{118} \right] + \left[\frac{0.57 \times 0.43}{127} \right]} = \pm 12.1\%$$

SAMPLE CALCULATIONS—RISK REDUCTION

	+ Outcome	- Outcome
+ Intervention	A 10	B 50
- Intervention	C 26	D 34

CER = Control Group Event Rate = $c / (c + d) = 0.43 = 43\%$

EER = Experimental Group Event Rate = $a / (a + b) = 0.17 = 17\%$

		Relative Risk Reduction (RRR)	Absolute Risk Reduction (ARR)	Number Needed to Treat (NNT)
CER	EER	$\frac{CER - EER}{CER}$	CER - EER	1/ARR
43%	17%	60%	26%	4
		*95% CI	10.3-41.7%	2-10

***95% Confidence Interval (CI) on an NNT = 1/(Limits on the CI of its ARR) =**

$$\pm 1.96 \sqrt{\left[\frac{CER \times (1 - CER)}{\# \text{ Control Pts}} \right] + \left[\frac{EER \times (1 - EER)}{\# \text{ Exper Pts}} \right]} =$$

$$\pm 1.96 \sqrt{\left[\frac{0.17 \times 0.83}{60} \right] + \left[\frac{0.43 \times 0.57}{60} \right]} = \pm 15.7\%$$

YOUR CALCULATIONS

	+ Outcome	- Outcome
+ Intervention	a	b
- Intervention	c	d

CER = Control Group Event Rate = $c / (c+d)$ =

EER = Experimental Group Event Rate = $a / (a+b)$ =

		Relative Benefit Increase (RBI)	Absolute Benefit Increase (ABI)	Number Needed to Treat (NNT)
EER	CER	$\frac{EER - CER}{CER}$	EER - CER	1/ABI
		*95% CI		

***95% Confidence Interval (CI) on an NNT =**
1/(Limits on the CI of its ABI) =

		Relative Risk Reduction (RRR)	Absolute Risk Reduction (ARR)	Number Needed to Treat (NNT)
CER	EER	$\frac{CER - EER}{CER}$	CER - EER	1/ARR
		*95% CI		

***95% Confidence Interval (CI) on an NNT =**
1/(Limits on the CI of its ABI) =

OUTCOMES—EVIDENCE APPRAISAL WORKSHEET

Citation:

Are the results of this outcomes measure study valid?

Appraisal Criterion	Comments
<p>Was this a study with more than one group?</p> <ul style="list-style-type: none"> • If not, describe the limitations resulting from the lack of a comparison group. 	
<p>Were the groups comparable at the start of the study?</p> <ul style="list-style-type: none"> • If not, briefly describe the differences and discuss the potential consequences for this study's results. 	
<p>If groups were not equal at the start of the study, was risk adjustment performed?</p> <ul style="list-style-type: none"> • If not, describe the potential consequences for this study's results. 	
<p>Were variables operationally defined and adequately measured by the data used for the study?</p> <ul style="list-style-type: none"> • If not, describe what proxy measures were used and the implications of their use for this study's results. 	
<p>Was a standardized outcomes instrument used?</p> <ul style="list-style-type: none"> • If not, describe how the outcomes were measured and the implications of this approach for the study's results. 	

Appraisal Criterion	Comments
Were standardized data collection methods implemented? • If not, describe how data were collected and the implications of this approach for the study's results.	
Did the intervention precede the outcome? • If not, describe the potential consequences for this study's results.	
Were other potentially confounding variables accounted for in the analysis? • If not, describe the potential consequences for this study's results.	
Were missing data dealt with in an appropriate manner? • If not, describe the potential consequences for this study's results.	

Are the valid results of this outcomes study important?

Appraisal Criterion	Comments
What were the statistical findings of this study?	<ul style="list-style-type: none"> • Tests of Differences? P-values? Confidence Intervals? • Effect Sizes? P-values? Confidence Intervals? • Absolute and Relative Benefit Increases? P-values? Confidence Intervals? • Absolute and Relative Risk Reductions? P-values? Confidence Intervals? • Number Needed to Treat (Harm)? Confidence Intervals? • Other?

Appraisal Criterion	Comments
What is the meaning (application) of these statistical findings for your patient/client case?	

Can you apply this valid, important evidence about outcomes in caring for your patient/client?

Appraisal Criterion	Comments
Are the study subjects similar to your patient/client? • If not, how different? Can you use this intervention in spite of these differences?	
Do the potential benefits outweigh the potential risks of using this intervention to achieve the specified outcome with your patient/client?	
Do the interventions and related outcomes fit within your patient/client's stated values or expectations? • If not, what will you do now?	

Additional notes:

SYSTEMATIC REVIEW— EVIDENCE APPRAISAL WORKSHEET

Citation:

Are the results of this systematic review valid?

Appraisal Criterion	Comments
Is this a systematic review of randomized trials? • If not, what types of studies are included? • What are the potential consequences of including these studies for this review's results?	
Does the review include a methods section that describes identifying and selecting all relevant trials? • If not, what are the potential consequences for this review's results?	
Does the review include a methods section that describes the processes and tools used to assess the quality of individual studies? • If not, what are the potential consequences for this review's results?	
Was the quality of individual studies reported? • If not, what are the potential consequences for this review's results?	
Was publication bias addressed? • If not, what are the potential consequences for this review's results?	
If this is a meta-analysis, were the individual patient data used in the analysis (or aggregate data)? • If not, what are the implications of using aggregated data?	

Are the valid results of this systematic review important?

Appraisal Criterion	Comments
Were the results consistent from study to study (i.e., homogeneous)? • If not, what are the potential consequences for this review’s results?	
If this paper is a meta-analysis, what are the statistical results (e.g., effect sizes, odds ratios, relative risks, likelihood ratios, numbers needed to treat) and associated confidence intervals?	
If this paper is not a meta-analysis, is there a substantive conclusion that can be drawn about the cumulative weight of the evidence (e.g., do the results all point in the same direction)?	

Can you apply this valid, important evidence from a systematic review in caring for your patient?

Appraisal Criterion	Comments
Is your patient different from those in the systematic review? • If yes, what are the important differences and what are their implications for your use of this study?	
Is the treatment feasible in your setting?	

Are your patient’s values and preferences satisfied by the regimen and its consequences?

If subgroup analysis was performed, should you believe apparent qualitative differences in the effectiveness of the intervention in some subgroups of patients? Only if you can say ‘yes’ to all of the following:

Appraisal Criterion	Comments
Do the qualitative differences really make biologic and clinical sense? • If not, what are the potential consequences for this review’s results?	
Is the qualitative difference both clinically (beneficial for some, but useless or harmful for others) and statistically significant? • If not, what are the potential consequences for this review’s results?	
Was this difference hypothesized before the study began (rather than the product of dredging the data), and has it been confirmed in other independent studies? • If not, what are the potential consequences for this review’s results?	
Was this one of just a few subgroup analyses carried out in this study? • If not, what are the potential consequences for this review’s results?	

Additional notes:

Alternative ways to evaluate the importance of statistical results from meta-analyses or individual papers using odds ratios.

Translating odds ratios to numbers needed to treat (NNTs):

The numbers in the body of the tables are the NNTs for the corresponding odds ratio at that particular patient’s expected event rate (PEER).

1. When the odds ratio (OR) < 1

This table applies when a bad outcome is prevented by therapy.

		OR < 1				
		0.9	0.8	0.7	0.6	0.5
Patient’s Expected Event Rate (PEER)	0.05	2.09 ^a	104	69	52	41 ^b
	0.10	110	54	36	27	21
	0.20	61	30	20	14	11
	0.30	46	22	14	10	8
	0.40	40	19	12	9	7
	0.50	38	18	11	8	6
	0.70	44	20	13	9	6
	0.90	101 ^c	46	27	18	12 ^d

^aThe Relative Risk Reduction (RRR) here is 10%.

^bThe RRR here is 49%.

^cThe RRR here is 1%.

^dThe RRR here is 9%.

2. When the odds ratio (OR) > 1

This table applies both when a good outcome is increased by therapy and when a side effect is caused by therapy.

		OR > 1				
		1.1	1.2	1.3	1.4	1.5
Patient’s Expected Event Rate (PEER)	0.05	212	106	71	54	43
	0.10	112	57	38	29	23
	0.20	64	33	22	17	14
	0.30	49	25	17	13	11
	0.40	43	23	16	12	10
	0.50	42	22	15	12	10
	0.70	51	27	19	15	13
	0.90	121	66	47	38	32

What are the potential benefits and harms to your patient from the therapy?

Method I: In the OR tables above, find the intersection of the closest odds ratio from the systematic review and your patient's expected event rate (PEER).

What are the potential benefits and harms to your patient from the therapy?

Method II: To calculate the NNT from any RR or OR and any PEER:

For RR < 1:

$$\text{NNT} = 1 / (1 - \text{RR}) \times \text{PEER}$$

For RR > 1:

$$\text{NNT} = 1 / (\text{RR} - 1) \times \text{PEER}$$

For OR < 1:

$$\text{NNT} = \frac{1 - [\text{PEER} \times (1 - \text{OR})]}{(1 - \text{PEER}) \times (\text{PEER}) \times (1 - \text{OR})}$$

For OR > 1:

$$\text{NNT} = \frac{1 + [\text{PEER} \times (\text{OR} - 1)]}{(1 - \text{PEER}) \times (\text{PEER}) \times (\text{OR} - 1)}$$

