Worksheet for Structured Review of Physical Exam or Diagnostic Test Study

Title of Manuscript:_____
Authors of Manuscript:_____
Journal and Citation:_____
Identify and State the Hypothesis
Primary Hypothesis:_____
Secondary Hypothesis:_____

Methods Questions:

1.) Does this Diagnosis Study use a Cross Sectional Study Design? ___Yes, ___No. If not what study design is used? _____
2.) What is the Diagnostic Test?_____
3.) What is the Gold-Standard?_____
3a.) Appropriate Gold Standard Used? ___Yes, ___No
3b.) Alternative Gold Standards_____
4.)What is the Target Disorder the Test is Trying to Diagnose?_____
5.) What is the Patient Population in whom the Test is Applied?_____
4a.) Appropriate Patient Population Studied?____Yes, ____No
4b.) If No, What is the Appropriate Patient Population?_____

Evaluation of Bias:

Were Examiners Blinded (Independent Examiner Used)? ___Yes, ___No

Are there sources of Selection Bias:

Filter Bias ____Yes ___No Explain:_____
Spectrum Bias? ___Yes, ____No Explain:_____
Subgroup Bias?___Yes, ____No Explain:_____
Verification Bias?____Yes, ____No Explain:_____
Incorporation Bias?____Yes, ____No Explain:_____

Are there sources of Observer Bias:

Absence of a Definitive Test? ___Yes, ____No Explain_____
Review or Interpretation Bias?____Yes, ____No Explain:_____
Context Bias?_____Yes, _____No Explain:_____

Are there Miscellaneous Sources of Bias

Could the test be graded as indeterminate? ___Yes ___No
Is the test result reader dependent?_____Yes, ____No
Has continuous test data been converted to discrete data to determine a significant threshold value? ___Yes, ___No
Is the test reported subject to future improvements (i.e. in technology) ___Yes, ___No
Is the study sponsored by the technology company ___Yes, ___No

Statistical Analysis

Were appropriate statistical methods used (e.g. type of test, estimator or confidence interval; choice of paired vs unpaired assessment; adjustment for baseline variables)?

□ Yes, and methods required minimal assumptions.
□ Yes, distributional (normal or non-normal) and other assumptions were justified.
□ Probably; methods are appropriate in general for the type of response variable, but assumptions were not justified.
□ No, but the test used is unlikely to affect the conclusions.
□ No, and there is a good possibility that the inappropriate statistical methods used have affected the conclusions.

Is Statistical Consultation Requested? □Yes □No Who should do it?_____

RESULTS
Enter Data Here:

	DISEASE +	DISEASE -	
TEST +	a=	b=	a+b=
TEST-	c=	d=	c+d=
	a+c=	b+d=	a+b+c+d=

D+ Disease Present D- Disease Absent
T+ Test Positive T- Test Negative

a=TP = True Positive b=FP=False Positive c=FN=False Negative d=TN=True Negative

SENSITIVITY = TPR= True Postive Rate = TP/(TP+FN) OR a/(a+c)_____
SPECIFICITY =TNR= True Negative Rate = TN/(FP+TN) OR d/(d+b)_____
ACCURACY = a+d/a+b+c+d _____
POSITIVE PREDICTIVE VALUE = a/a+b _____
NEGATIVE PREDICTIVE VALUE = d/c+d _____

FPR=False Positive Rate=FP/(FP+TN) OR b/(b+d) _____
FNR = False Negative Rate = FN/(TP+FN) OR c/(c+a)_____

LR(+)=Likelihood Ratio of a Positive Test = L(T+if D+)/L(T+ if D-)
or = TPR/FPR or = Sensitivty/FPR or = sensitivity/(1-specificity) _____

LR(-)=Likelihood Ratio of a Negative Test = L(T- if D+)/L(T- if D-)
or = FNR/TNR or = FNR/Specificity or = (1-sensitivity)/specificity _____

Pretest Probability of Disease=(Total D+/Total Population)=_____

Pretest Odds = P/(1-P) = _____

Post-Test Odds = P/(1-P) x LR+ or LR - _____

Post-Test Probability of Disease – O/(1+O) _____

SUMMARY OF DATA

Test	Gold Standard	N	Sens	Spec	Accuracy	PPV	NPV	LR +	LR -	Pretest Probability	Pretest Odds	Post-Test Odds	Post Test Prob

Are the Results Valid?
Was there an independent, blind comparison to a referecne (gold) standard of diagnosis? □Yes □No
Was the study test described adequately? □Yes □No
Was the diagnosis test evaluated in an appropriate specturm of patients? □Yes □No
Was the reference standard applied regardless of the diagnostic test result? □Yes □No
Has the utility of the test been determined? □Yes □No
What is the impact of the results? □Yes □No

Can the Results be Applied to My Patients?
Is the diagnostic test available, affordable, accurate, and precise in my setting? □Yes □No
Can I come up with a reasonable pre-test probability of the disease for my patient? □Yes □No
Will the post-test probability change the management of my patient? □Yes □No

APPENDIX:

LEVELS OF EVIDENCE FOR PRIMARY RESEARCH QUESTION

	Types of Studies			
	Therapeutic Studies—Investigating the Results of Treatment	Prognostic Studies—Investigating the Outcome of Disease	**Diagnostic Studies—Investigating a Diagnostic Test**	Economic and Decision Analyses—Developing an Economic or Decision Model
Level I	1. Randomized controlled trial a. Significant difference b. No significant difference but narrow confidence intervals 2. Systematic review[2] of Level-I randomized controlled trials (studies were homogeneous)	1. Prospective study[1] 2. Systematic review[2] of Level-I studies	**1. Testing of previously developed diagnostic criteria in series of consecutive patients (with universally applied reference "gold" standard)** **2. Systematic review[2] of Level-I studies**	1. Clinically sensible costs and alternatives; values obtained from many studies; multiway sensitivity analyses 2. Systematic review[2] of Level-I studies
Level II	1. Prospective cohort study[3] 2. Poor-quality randomized controlled trial (e.g., <80% follow-up) 3. Systematic review[2] a. Level-II studies b. Nonhomogeneous Level-I studies	1. Retrospective study[4] 2. Study of untreated controls from a previous randomized controlled trial 3. Systematic review[2] of Level-II studies	**1. Development of diagnostic criteria on basis of consecutive patients(with universally applied reference "gold" standard)** **2. Systematic review[2] of Level-II studies**	1. Clinically sensible costs and alternatives; values obtained from limited studies; multiway sensitivity analyses 2. Systematic review[2] of Level-II studies
Level III	1. Case-control study[5] 2. Retrospective cohort study[4] 3. Systematic review[2] of Level-III studies		**1. Study of nonconsecutive patients (no consistently applied reference "gold" standard)** **2. Systematic review[2] of Level-III studies**	1. Limited alternatives and costs; poor estimates 2. Systematic review[2] of Level-III studies
Level IV	Case series (no, or historical, control group)	Case series	**1. Case-control study** **2. Poor reference standard**	No sensitivity analyses
Level V	Expert opinion	Expert opinion	**Expert opinion**	Expert opinion

1. All patients were enrolled at the same point in their disease course (inception cohort) with greater than or equal to 80% follow-up of enrolled patients.
2. A study of results from two or more previous studies.
3. Patients were compared with a control group of patients treated at the same time and institution.
4. The study was initiated after treatment was performed.

5. Patients with a particular outcome ("cases" with, for example, a failed total arthroplasty) were compared with those who did not have the outcome ("controls" with, for example, a total hip arthroplasty that did not fail).

Definitions

Sources of Selection Bias:

Filter Bias: Not all patients have characteristics that would lead one to order a diagnostic test. Ask: Were only a portion of the eligible patients tested? If patients are inappropriately filtered before entry into the study, the test may not have external validity.
Spectrum and subgroup Bias: Occurs when a test is more accurate for advanced stages of a disease, but not for early stages. What stage was used in this experiment?
Verification Bias: Occurs when patients are selected to reciever the gold-standard test based upon the results of the diagnostic test being evaluated. All subjects (postive and negative for test under evaluation) should receive the gold standard test.
Incorporation Bias: Occurs then the diagnostic test being studied is used as or is part of the gold standard. The diagnostic test and the gold standard should be independent of each other.

Sources of Observer Bias:

Absence of a Definitive Test: Occurs when no gold standard exists or the gold standard is poor quality.
Review or Interpretation Bias: Occurs when investigators are not blinded to the nature of the patient's other test results. Will artificially inflate the utility of the test under study. Test Review Bias occurs when the person interpreting the test has knowledge of the patient's gold standard test result. Diagnostic Review Biasl occurs when the person interpreting the gold standard test knows the result of the diagnostic test.
Context Bias: Occurs then the person interpreting the test bases the reading upon known clinical information.

Miscellaneous Sources of Bias

Indeterminate and Uninterpretable Results: When a test is not clearly positive or negative and requires interpretation.
Reproducibility Problems: Occurs when the performance of a diagnostic test depends upon the performance of the technician and the equipment used in the performance of the test.
Post-hoc Selection of Test Positivity Criteria: Occurs when the test results are continuous data, but the researchers assign the data as catagoriacl with a specific point for a positive or negative test. Post-hoc data dredging/analysis could introduce a Type I error. Solved by using likelihood ratios ans sensitivity and specificity.
Temporal Changes: Test characteristics measured at one point in time may change as the test is technically improved. Measures calculated from the newer technology do not apply to the older technology.
Publication Bias/Conflicts of Interest: Studies that are not positive rarely get published, especially if there is a conflict of interest and the study is sponsored by industry.

Formulas:

D+ Disease Present
D- Disease Absent
T+ Test Positive
T- Test Negative

	D+	D-
T+	**TP**	**FP**
T-	**FN**	**TN**

TP = True Positive
FP=False Positive
FN=False Negative
TN=True Negative

TPR= True Postive Rate = TP/(TP+FN) AKA Sensitivity
TNR= True Negative Rate = TN/(FP+TN) AKA Specificity
FPR=False Positive Rate=FP/(FP+TN)
FNR = False Negative Rate = FN/(TP+FN)
LR(+)=Likelihood Ratio of a Positive Test = L(T+if D+)/L(T+ if D-)
or = TPR/FPR or = Sensitivty/FPR or = sensitivity/(1-specificity)
LR(-)=Likelihood Ratio of a Negative Test = L(T- if D+)/L(T- if D-)
or = FNR/TNR or = FNR/Specificity or = (1-sensitivity)/specificity