The 3-min appraisal of a diagnostic test

Teresa Chien, Rajesh Malhotra[1], Mohit Bhandari

Introduction

Diagnostic tests are invaluable tools used in various healthcare settings to distinguish between patients who have a disease and those who do not. It is essential for surgeons to be skilled in critically appraising published papers about a diagnostic test.[1] This article will provide some simple and quick guidelines to assist in the 3-min critical appraisal process of the literature on diagnostic tests.

Key Criteria for Critical Appraisal

Many aspects of diagnostics need to be evaluated, and there are three specific areas that should be critically appraised in diagnostic test studies as described by Guyatt in his User's Guide to the Medical Literature: validity of the study, results, and the applicability of the diagnostic test [Table 1].

Are the results of the study valid?

The validity of a diagnostic test study can be critically appraised through examining the study design. The patient population of the study should include a wide spectrum of patients with varying disease conditions and stages of treatment to ensure that there is genuine diagnostic uncertainty.[2,3] Diagnostic uncertainty increases when symptoms of the target condition are also characteristic of other diseases.[3] To minimize misjudgment of the study results, there should be a variety of patients.

Another essential component to analyze is if an independent (blind) comparison between the diagnostic test and an appropriate reference standard was done for each patient. For example, a surgeon wishing to understand whether the "Lachman" test for cruciate ligament is predictive of a tear could confirm the accuracy of the test by ensuring all patients tested received an independent "standard" confirmatory test (magnetic resonance imaging (MRI) or arthroscopic visualization). To increase the study validity and minimize potential bias in overestimating test outcomes, those interpreting the test results should be blinded and be different than those interpreting the reference standard.[3] In this case, the surgeon performing the physical exam manoeuvre and grading the presence/absence of a positive Lachman would not be the same surgeon who reviewed the gold standard confirmatory test (MRI or arthroscopy).

The results of a diagnostic test should not influence a decision to perform the reference standard. This is referred to as verification bias.[3] For example, verification bias results when only patients with a positive "Lachman" test get an MRI (or arthroscopy), but those with a negative test do not. A good study design should take verification bias into account and take measures for bias prevention.

What are the results?

The terms sensitivity and specificity are often used to describe the effectiveness of a test, but the literature has

Table 1: Guidelines to critically appraising literature on diagnostic tests

Are the results of the study valid?[a]
- Is the patient population suitable to ensure there is genuine diagnostic uncertainty?
- Is there an independent (blind) comparison with an appropriate reference standard for every patient?
- Did the author clearly define the diagnostic test and the reference standard?
- Is there a good study design to prevent verification bias?

What are the results?
- Are likelihood ratios being used as a tool to aid in clinical decision making?
- Did the authors present specificity, sensitivity, pretest and posttest probabilities to effectively apply likelihood ratios?
- Are the results of this study useful in my practise?
- Does the diagnostic test study relate to the needs of my patient population?

Will my patient benefit from having this test?
- Can the diagnostic test results be easily reproduced in patients of my practise?

[a]From Guyatt and coworkers.[2]

McMaster University, Hamilton, Canada, [1]Department of Orthopaedics, AIIMS, Delhi, India

Address for correspondence: Dr. Mohit Bhandari,
2309 Hoover Court, Hamilton, L7P 4V2, ON, Canada. E-mail: bhandam@mcmaster.ca

Access this article online

Quick Response Code:

Website: www.ijoonline.com

DOI: 10.4103/0019-5413.80317

shown that likelihood ratios are a better statistical tool in aiding clinical decision making.[2] Sensitivity is the proportion of individuals with a target condition that test positive and specificity is the proportion of individuals without a target condition that test negative. These dichotomous measures indicate either positive or negative tests, while likelihood ratios account for the cases in the middle of a wide spectrum of patients.[2,3]

Likelihood ratios combine specificity and sensitivity, which provides a better measure of the test as it links the pretest probability (prevalence of disease) to the post-test probability (chances of detecting the target condition with the diagnostic test).[3,4] Likelihood ratios are a ratio of the proportion of positive test results versus the proportion of negative test results.[5] A likelihood ratio equal or close to 1 means that the test has minimal value because it cannot differentiate between those who have the target condition and those who do not.[4] A large likelihood ratio (>1) means a larger proportion of the test results will occur in positive patients, whereas a smaller ratio (<1) indicates a higher probability that the test result will occur more frequently in healthy patients.[5] The ability to understand and interpret these ratios is essential in understanding quantitative features of a study.

Are the results of this study useful in my practise?

The last component to consider is the applicability of the study results to current practise. It is crucial to compare one's patient to the patient population in the study. The greater the similarities, the more appropriate and relevant the study is to one's practise. An accurate diagnostic test with minimal risks to the patient can be an invaluable stepping stone to improving healthcare. A diagnostic test that is easy to conduct and reproduce allows for a better integration into clinical practise.[3]

A Practical Example

In the study by Iannotti *et al.*, an office-based ultrasonography was examined for its accuracy in diagnosing rotator cuff tears.[6]

Are the results of the study valid?

The patient population of Iannotti *et al.*'s study were all clinically diagnosed with rotator cuff symptoms. By having this inclusion criterion, it limits a wide spectrum of patients which may reduce the study validity.

This study clearly defined the diagnostic test (ultrasonography) and compared it with an appropriate reference standard (operative findings); however, some areas of the study design allowed for potential bias. The study design was complex involving consecutive radiographs, clinical examinations, ultrasounds, and preoperative MRI scans prior to the study. The orthopedic surgeon did the initial physical examinations and interpreted the radiographs. The surgeon's involvement in the preliminary assessment of the patient might have caused bias to his/her diagnosis after evaluating the ultrasound and preoperative MRI results. The authors' rationale behind their design was that the diagnostic study should be carried out in a similar way similar to a typical clinical practise, where patients will undergo similar diagnostic procedures. The authors also suggested that the surgeon's bias from a previous involvement with the patient might actually lead to an increased accuracy of the test then the blinded radiologist.

What are the results?

The ultrasound had an accuracy of 80%, sensitivity of 88%, positive predictive value of 79%, negative predictive value of 90%, and a false-positive rate at 21%. These results were reported in a 3 × 3 table comparing three categories: no tear, partial-thickness tear only, and full-thickness tear with or without partial-thickness tear. If the latter two categories were combined to represent the target condition, and configured into a 2 × 2 contingency table, the likelihood ratios can be calculated.

The process of calculating likelihood ratios are shown in Table 2. The positive likelihood ratio is 4.8 and the negative likelihood ratio is 0.05. This indicates that a positive ultrasound is 4.8 times more like to occur in patients with a rotator cuff tear (either partial or full-thickness). Likelihood ratios are better measures of interpreting the results.

Are the results of this study useful in my practise?

While this study had some limitations in its design, it concluded that office-based ultrasounds can be used effectively to diagnose rotator cuff tears when there are well-trained staff, preliminary clinical examinations, and radiographs.

Summary

The ability to critically appraise the literature on diagnostic tests is essential for interpreting and understanding results of a study. The validity of the study, the results, and

Table 2: Likelihood ratio calculations from the practical example

Ultrasonographic findings	Operative findings		
	Tears present	No tears	Total
Tears present	True positive (A) 76	False positive (B) 4	80
No tears	False negative (C) 3	True negative (D) 16	19
Total	79	20	99

Sensitivity: A/(A + C) = 76/(76 + 3) = 0.96, Specificity: D/(B + D) = 16/(4+16) = 0.80
Positive likelihood ratio = sensitivity/(1 –specificity) = 0.96/(1 –0.80) = 4.8
Negative likelihood ratio = (1 – sensitivity/specificity) = (1 – 0.96)/0.80 = 0.05

applicability of results to your patient population must be considered appropriately before applying the new knowledge in evidence-based orthopedics.

References

1. Mundi R, Chaudhry H, Singh I, Bhandari M. Checklists to improve the quality of the orthopaedic literature. Indian J Orthop 2008;42:150-64.
2. Greenhalgh T. How to read a paper: Papers that report diagnostic or screening tests. BMJ 1997;315:540-3.
3. Bhandari M, Montori VM, Swiontkowski MF, Guyatt GH. User's guide to the surgical literature: How to use an article about a diagnostic test. J Bone Joint Surg Am 2003;85:1133-40.
4. Dujardin B, Van den Ende J, Van Gompel A, Ugner JP, Van der Stuyft P. Likelihood ratios: A real improvement for clinical decision making? Eur J Epidemiol 1994;10:29-36.
5. Richardson WS, Wilson MC, Keitz SA, Wyer PC; EBM Teaching Scripts Working Group. Tips for Teachers of evidence-based medicine: Making sense of diagnostic test results using likelihood ratios. J Gen Intern Med 2008;23:87-92.
6. Iannotti JP, Ciccone J, Buss D, Visotsky JL, Mascha E, Cotman K, *et al.* Acuracy of office-based ultrasonography of the shoulder for the diagnosis of rotator cuff tears. J Bone Joint Surg Am 2005;87:1305-11.

How to cite this article: Chien T, Malhotra R, Bhandari M. The 3-min appraisal of a diagnostic test. Indian J Orthop 2011;45:389-91.

Source of Support: Nil, **Conflict of Interest:** None.