Introduction to Diagnostic Reasoning

by Michael Turlik, DPM[1] ✉

This is the first of four articles written for podiatric physicians to help them understand and apply the results of diagnostic studies to their practice. This article deals with how clinicians arrive at a diagnosis and how to interpret results from a diagnostic trial. An article from the foot and ankle literature will be used to illustrate the concepts discussed in the publication.

The method of arriving at a diagnosis can be a simple or a very complex process. This depends upon the clinician's knowledge, experience, clinical presentation of the diagnostic problem, prevalence of the disease and diagnostic studies employed. Podiatric physicians always encounter some degree of uncertainty in practice whether it is the true effect of a therapeutic intervention or the diagnosis of a patient's condition. After collecting information needed to make a diagnosis there usually is some information threshold after which additional information becomes irrelevant and treatment begins. (Fig.1) There are two basic ways by which clinicians arrive at a diagnosis. They are, pattern recognition / categorization or a probabilistic diagnostic reasoning / hypotheico-deductive approach.[1,2]

Pattern recognition approach or Categorization

This approach is used by experts making a common diagnosis in their field of expertise. The use of this method varies widely among clinicians and is based upon the thoroughness of their knowledge base and their experience. When using this approach podiatric physicians are able to quickly evaluate the clinical scenario and place it in some familiar combination of signs and symptoms and rapidly make the diagnosis. This type of diagnostic reasoning does not involve the generation of multiple hypotheses which are tested and it is unlikely they use the same reasoning process as novice clinicians. For example, Mr. Jones a 52 year-old obese white male who presents to the office complaining of a two-week history of heel pain which began insidiously and is localized to the plantar medial aspect of his heel.

Address correspondence to: Michael Turlik, DPM
Email: mat@evidencebasedpodiatricmedicine.com

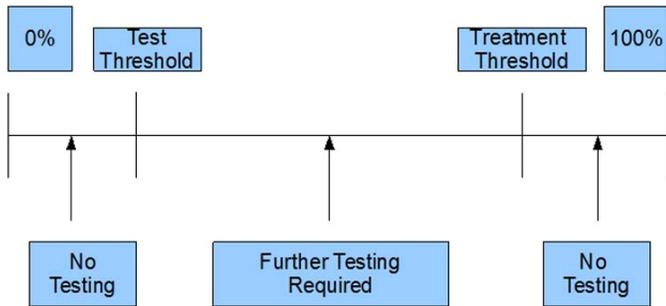[1] Private practice, Macedonia, Ohio.

Figure 1 Threshold model of decision making.*

*Reproduced with permission from Center for Evidence-Based Medicine,
http://www.cebm.net/index.aspx?o=1043

He relates the pain is worse after periods of inactivity, arising from a sitting position or when first bearing weight on the heel after sleeping. Physical examination reveals no redness, no edema, no deformity, but tenderness to palpation of the plantar medial heel. Even the inexperienced podiatrist would be able to make the diagnosis of mechanically induced heel pain given this scenario. This diagnosis does not require any type of diagnostic studies for most podiatric physicians and the diagnosis is thought to be clinical in nature.[3] The pretest probability of mechanically induced heel pain in this scenario is very high and likely to be greater than 90%. In addition, the podiatrist will recognize that the chance of a bone tumor of the calcaneus producing this clinical picture is close to 0%. The percentage of patients with the disease in the specified population at any point in time is defined as prevelance / pretest probability. Pretest probabilities which are extremely low or extremely high usually will not benefit from further diagnostic testing. (Fig. 1)

Probabilistic Diagnostic Reasoning or Hypotheico-Deductive

When clinicians face an atypical presentation of a common condition or something more challenging for their specialty, clinicians will switch from pattern recognition to probabilistic diagnostic reasoning.

As a result of the clinical encounter the clinician will generate a short list of diagnostic hypotheses with an estimate of the probability for each possibility. This list will guide subsequent efforts in data collection. The pretest probability for this type of diagnostic inquiry usually lies in an intermediate range rather than at the extremes. (Fig. 1) Therefore, diagnostic studies may be very helpful in distinguishing between the different hypotheses and result in restructuring and reprioritizing diagnostic possibilities as further information is obtained. For example, a 50-year-old neuropathic diabetic male presents with a one week history of progressive redness, swelling and pain about a recurrent plantar ulcer successfully treated with oral antibiotics and local wound care in the past. Physical examination reveals a mildly obese afebrile male with palpable pulses and lack of protective sensation bilaterally. The ulcer under the first metatarsophalangeal joint (MPJ) measures 1.5 cm in diameter and exhibits a red base. There is minimal drainage on the dressing without odor. The important question which needs to be answered in this scenario is; does this patient have osteomyelitis? Pretest probability (prevalence) in this case varies from 20-66% based upon the study location referenced.[4] Higher pre-test probabilities are seen in tertiary care hospital settings, lower in outpatient primary care settings. The range of pretest probabilities in this case differs from the earlier example of mechanically induced heel pain because it is in the intermediate category indicating that some further diagnostic test(s) is (are) necessary. (Fig. 1)

The test and treatment thresholds (Fig. 1) are not static but dynamic. They vary with the invasiveness and cost of the test, the consequences of misdiagnoses of the disease process, and the efficacy and expense of the treatment. (Table 1) For example, in the case of a diabetic patient with a pedal ulcer referenced above, the test threshold would be lower for using a metal probe to evaluate the ulcer for osteomyelitis than for performing a bone biopsy. Since mechanically induced heel pain is a benign self limited condition which responds to non surgical care the treatment threshold would be lower than the treatment threshold for osteomyelitis in a diabetic patient. A comprehensive explanation of how to calculate test treatment thresholds using decision tree analysis is provided for the interested reader.[5]

	Test threshold	Treatment threshold
Increased	More invasive and expensive the test strategy, less serious the condition if undiagnosed, unsure of treatment options.	Less invasive and expensive the test, the more benign the prognosis, increased costs and adverse effects of treatment options.
Decreased	Safer and less costly test, more serious the condition if left undiagnosed, safer and more available treatment options.	More invasive the test, the poorer the prognosis, the safer and less costly the proposed treatment.

Table 1 Variations in test / treatment threshold.[2]

	Disease Present	Disease Absent
Test Positive	True positives (TP)	False positives (FP)
Test Negative	False positives (FP)	True negatives (TN)

Table 2   2 x 2 diagnostic table.

The information gained from the results of a diagnostic study change the pretest probability, the revised estimate of prevalence is termed the posttest probability. The magnitude of the change is a function of the strength of the diagnostic intervention on the pretest probability. The direction of the posttest probability can either be higher or lower than the pretest probability depending upon the results of the diagnostic study used. The strength of the diagnostic intervention may be presented in many ways; the most clinically useful is a likelihood ratio.

Assessing the Performance of a Diagnostic Test

The question that podiatric physicians must answer after ordering a diagnostic test is; based upon the results of this test how probable is it that my patient has a diagnosis of ___? To answer this question it is necessary to construct a 2 x 2 table (Table 2) from a study of the intervention to determine the strength of the test. Some measures of probability derived from a 2 x 2 table are the following:

Sensitivity: the proportion of the patients with the disease who test positive.
TP / (TP + FN)

Specificity: the proportion of the patients without the disease who test negative.
TN / (TN+FP)

Positive Predictive Value: proportion of patients with a positive test who have the disease
TP / (TP+FP)

Negative Predictive Value: proportion of patients with negative test who do not have the disease
TN / (TN+FN)

Positive Likelihood Ratio: how much the odds of the disease increase when a test is positive.
sensitivity / 1-specificity

Negative Likelihood Ratio: how much the odds of the disease decrease when a test is negative.
1-sensitivity / specificity

The higher the sensitivity of the test is the better its ability to detect disease due to a low false negative rate. Diagnostic tests with a high sensitivity 95-99% are used when there is an important price for missing a serious disease which is treatable. High sensitivity tests are usually used early in the workup of the disease and if positive are followed up with a test which has a high specificity. If a test with a high sensitivity is negative the podiatric physician can be comfortable in ruling out the disease process. The mnemonic **SnOut** refers to a diagnostic test with a high sensitivity.

Diagnostic tests which have a high specificity are used to identify those patients who do not have the condition of interest. A highly specific test will rarely miss people as having the disease when they do not. These types of tests are most useful to confirm a diagnosis which has been suggested by a test which is highly sensitive. Highly specific tests are particularly useful when false positive results can cause harm to the patient physically, psychologically, or fiscally. If the test results are positive it is very helpful to the podiatric physician in confirming the disease process. The mnemonic **Spin** refers to a diagnostic test with a high specificity.

It is not possible to have a test which is highly specific and sensitive when dealing with data collected over a range of values.

When the test is measured over a continuum of values changing the artificial cutoff point causes a change in the sensitivity and specificity. Sensitivity can only be increased at the expense of specificity. Sensitivity and specificity are not clinically useful measures and do not answer the question of probability of having or not having the disease under evaluation.[6]

Predictive values are another measurement of test efficiency which can be derived from a 2 x 2 table.[7] predictive values derived can be used to gain information regarding the probability of disease in patients. As a test's sensitivity increases so too does the negative predictive value. As a test's specificity increases likewise positive predictive values increase. *Unlike sensitivity and specificity predictive values are influenced by disease prevalence.* Predictive values vary with disease prevalence in a nonlinear manner.[8] Therefore using predictive values derived in outpatient primary care setting will be misleading when applied to a tertiary care setting since the prevalence is usually different. This is a major limitation to the podiatric physician when using predictive values in clinical practice. In order to be clinically useful they should be employed in as similar a practice setting they were derived from.

A third method of determining test efficiency from a 2 x 2 table is to generate likelihood ratios.[6] Likelihood ratios are not apt to be influenced by disease prevalence provided disease spectrum remains the same for a different prevalence.[9] Likelihood ratios are expressed as odds rather than proportions. Sensitivity, specificity, as well as predictive values are expressed as proportions. Likelihood ratios are the preferred method of expressing test efficiency in evidence-based medicine publications. Likelihood ratios combine sensitivity and specificity of a diagnostic study which allows you to intuitively determine the odds of which the pretest probability will change based upon a positive or negative test result. Pretest probability X likelihood ratio = posttest probability. Using likelihood ratios to modify pretest probabilities to determine the posttest probability cannot be done directly.

	LIKELIHOOD RATIO	95% CI*
ULCER AREA > 2 cm^2	7.2	1.1-49
ABNORMAL RADIOGRAPH	2.3	1.6-3.3
POSITIVE MRI	3.8	2.5-5.8
POSITIVE PROBE TO BONE	6.4	3.6-11
ESR > 70 mm/hr	11	1.6-79

Table 3 Likelihood Ratios for Studies used to Evaluate Diabetic Osteomyelitis.[11] (*Confidence Intervals)

Since likelihood ratios are expressed as odds rather than proportions they must be converted prior to application. This can be done using mathematical conversions, internet calculators, or a nomogram.

How best to estimate the prevalence of disease? Clinical observations and experience are often inaccurate. A better estimate arises from reviewing the medical literature on the subject and/or evaluation of large computerized databases. Pretest probability is not a constant but varies with the clinical environment. Prevalence is increased when patients are passed through a filter from a primary care source to a tertiary care facility.

In order for a podiatric physician to correctly utilize a diagnostic study he or she will need to estimate the prevalence of the disease in their patient population, the likelihood ratio of the test employed and the rigor of the study used to determine the test's accuracy. In a recent systematic review of electrodignostic techniques currently in use to evaluate tarsal tunnel syndrome (TTS)[10] the authors conclude that due to the poor quality of the studies sensitivities and specificities reported could not be combined in a summary statistic. In addition, prevalence for TTS could not be determined. The author's conclusions limit the usefulness of electrodiagnostic studies in the evaluation of TTS.

Diabetes and Pedal Osteomyelitis

A recent article[11] appraises the published literature concerning the various diagnostic options for evaluating infected diabetic foot ulcers for the presence of osteomyelitis. The gold standard in each study was bone biopsy. A summary of the authors findings limited to higher quality studies is presented in Table 3. The highest likelihood ratios are found for ulcer area > 2 cm^2 and erythrocyte sedimation rate (ESR) > 700 mm/hr. Unfortunately, these test also have very large 95% confidence intervals which indicates that the results of these studies are not very precise. Tests which have a narrower 95% confidence interval are magnetic resonance imaging (MRI), probe to bone and abnormal radiograph. The probe to bone test based upon its cost, and adverse effects should be the first test under taken by the podiatric physician when evaluating an infected diabetic pedal ulcer for the presence of osteomyelitis.

The likelihood ratio for the probe to bone test cited in Butalia's review is a composite of three different studies. One of which is Lavery's study. Lavery and colleagues evaluated the accuracy of a probe to bone test for osteomyelitis in patients with diabetic foot ulcers.[4] They expressed their results in terms of sensitivity, specificity, positive and negative predictive values. They did not report likelihood ratios.

	Disease positive	Disease negative	Totals
Positive probe to bone test	26	20	46
Negative probe to bone test	4	197	201
Totals	30	217	247

Table 4 Results of probe to bone test.
*modified from Diabetes Care 30: 270, 2007

Level	Criteria
1	An independent, masked comparison with reference standard among an appropriate population of consecutive patients.
2	An independent, masked comparison with reference standard among nonconsecutive patients or confined to a narrow population of study patients.
3	An independent, masked comparison with an appropriate population of patients, but reference standard not applied to all study patients.
4	Reference standard not applied independently or masked.
5	Expert opinion with no explicit critical appraisal, based on physiology, bench research.

Table 5 Levels of evidence for diagnostic studies.[7]

In the results section the authors report information which could be used to construct a 2 x 2 table (Table 4). Using an online diagnostic calculator[12] likelihood ratios can be calculated. The value for a positive test was 9.4 and a negative test 0.14. When likelihood ratios are greater than one this increases the chance of the disease being present, likelihood ratios less than one decrease the chance of the disease being present. Likelihood ratios of > 10 or < 0.1 generate large conclusive changes. Likelihood ratios between 5-10 and 0.1-0.2 are associated with moderate changes in probability. The likelihood ratios calculated from Lavery's study[4] are associated with moderate/large changes in diagnostic probabilities. The pretest probability (prevalence) from Lavery's study[4] is 12%.

Using an on line calculator[13] or a Likelihood Ratio Nomogram (Fig. 2) the posttest probability can be calculated for a positive test to be 56.4% and a negative test 1.87%. A negative test should fall below the test threshold therefore effectively ruling out the condition. (Fig. 1) A positive test in this scenario still remains in the intermediate range for this prevalence and indicates further testing is necessary. If the prevalence were higher for example, 60% which is the prevalence seen in some studies in tertiary care centers[14] the posttest probability for a negative result would be 17.4% and a positive result would be 93.4%.
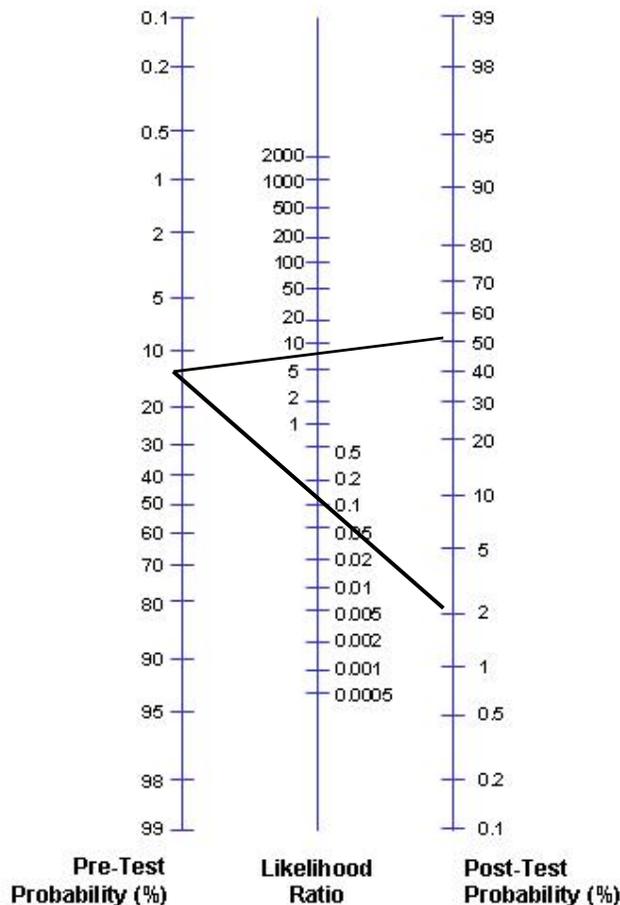
Figure 2 Likelihood Ratio Nomogram.*

*reproduced with permission from Center for Evidence-Based Medicine
 http://www.cebm.net/index.aspx?o=1043

These results indicate that further testing for a positive test is likely unnecessary while a negative test may fall within the intermediate range and may require further testing, opposite the results using the prevalence from Lavery's study.

The above example demonstrates the use of likelihood ratios for diagnostic studies evaluating a dichotomous outcome. Likelihood ratios can also be used with continuous test results as interval likelihood ratios.[15]

How believable is the likelihood ratio derived from a study?

The quality of the evidence derived from a diagnostic study is a function of the studies ability to minimize bias.[16] The best study design for diagnostic tests (Level 1) is an independent, masked comparison with a reference standard among an appropriate population of consecutive patients. Just as with randomized controlled trials, diagnostic studies are separated into different levels of evidence (Table 5), with the less rigorous (more biased) studies over estimating test effectiveness.[17] The largest effect of overestimation occurs from studies which include non-representative patients or studies which apply different reference standards for positive and negative test results. The smallest overestimation occurred when blinding was not adhered to during the study. The following article in this series will discuss how to critically appraise a diagnostic study for validity.

References

1. Elstein A, Schwartz:. Clinical problem solving and diagnostic decision making: a selective review of the cognitive research literature. In: Knottnerus JA (Ed). The Evidence Base of Clinical Diagnosis. London, England: BMJ books, 179 – 195, 2002.
2. Richardson WS, Wilson M: The process of diagnosis. In: Guyatt G, Bhandari M, Tornetta P, Schemitsch EH, Sprint Study Group: Users guides to the medical literature. New York, New York: McGraw-Hill, 399 – 406, 2008.
3. Cole C, ,Seto C, Gazewood J: Plantar fasciitis: Evidence-based review of diagnosis and therapy. Am Fam Physician 72: 2237 – 2242, 2005.
4. Lavery L, Armstrong DG, Peters EJG, Lipsky BA: Probe-to-Bone Test for Diagnosing Diabetic Foot Osteomyelitis. Diabetes Care 30: 270 – 274, 2007.
5. Pauker S, Kassirer J: The threshold approach to clinical decision making. NEJM 302: 1190 – 1116, 1980.
6. Deeks J, Altman D: Diagnostic tests 4: likelihood ratios. BMJ 329: 168 – 169, 2004.
7. Altman D, Bland JM: Statistics notes: Diagnostic tests 2: predictive values. BMJ 309: 102, 1994.
8. Predictive values http://www.poems.msu.edu/InfoMastery/Diagnosis/Predictive Values.htm. Accessed 8/25/2009. Accessed 09/09/2009.
9. Montori V, Wyer P, Newman T, Keitz S, Guyatt G: Tips for learners of evidence-based medicine: 5. The effect of spectrum of disease on the performance of diagnostic tests. CMAJ 173: 385 – 390, 2005.

10. Patel A, Gaines K., Malmut R., Park T, Del Toro D, Holland N: Usefulness of electrodiagnostic techiques in the evaluation of suspected tarsal tunnel syndrome: An evidence-based review. Muscle and Nerve 32: 236 – 240, 2005.

11. Butalia S, Palda V, Sargeant R, Detsky A, Mourad O: Does this patient with diabetes have osteomyelitis of the lower extremity? JAMA 299: 806 – 813, 2008.

12.   Likelihood Ratio Calculator http://araw.mede.uic.edu/cgi-alansz/testcalc.pl Accessed 3/8/2009.

13.   Post-test probability of disease calculator. http://homepage.mac.com/aaolmos/Posttest/posttest.html Accessed 3/9/2009.

14. Grayson ML, Gibbons GW, Balogh K, Levin E, Karchmer AW: Probing to bone in infected pedal ulcers. A clinical sign of underlying osteomyelitis in diabetic patients. JAMA 273: 721 – 723, 1995.

15. Mayer D: *Essential Evidence-based Medicine.* Cambridge, England: Cambridge University press, 233 – 236, 2004.

16. Moore A, McQuay H: Systematic reviews of diagnostic tests. In: *Bandolier's Little Book of Making Sense of the Medical Evidence.* London, England: Oxford University press, 236 – 242, 2006.

17. Lijmer JG, Mol BW, Heisterkamp S, Bonsel GJ, Prins MH, van der Meulen JHP, Bossuyt PMM: Empirical evidence of design-related bias in studies of diagnostic tests. JAMA 282: 1061 1066, 1999.