

Test-indication Curves

JOSEPH BERNSTEIN, MD

Test-indication curves (TICs) are tools for determining whether a test is indicated for a given patient. They apply the threshold approach of Pauker and Kassirer in graphic form. These curves are composed of two parts: the raw curve, which plots posttest probability versus pretest probability (given values for specificity and sensitivity); and the final curve, in which three straight lines are added to the raw curve by the clinician to generate a TIC for a given treatment threshold. In the final curve, the complete range of pretest probability is segregated into three zones, corresponding to the three groups described by Pauker and Kassirer: those patients in whom disease is assumed to be present and who are thus best treated empirically; at the other extreme, those who require neither testing nor treatment; and, finally, those in the middle, for whom the test is indicated, since the decision to treat would be based on the test result. Thus the clinician could consult the TIC and determine with certainty whether the test should be employed for a given patient. It also could be modified with ease for a different patient, with a different set of threshold values. TICs provide a complete, visual interpretation of a test's diagnostic power, in the context of a given treatment threshold. They foster an intuitive comprehension of Pauker and Kassirer's method, and offer the clinician a facile means to prove that a test is indicated in a given setting. By promoting the use of exactly those tests that are indicated, TICs can help spare the patient the cost, burden, and risk of unnecessary testing, and help spare the physician the cost, burden, and risk of interpreting inconclusive test results. *Key words:* test-indication curves; treatment thresholds; laboratory tests. (*Med Decis Making* 1997;17:103–106)

Test indication curves [TICs] are graphic tools for determining whether a diagnostic test is indicated in a given clinical setting. The correct indication for a test is, of course, that its result could alter the patient's clinical management: that the test might increase (or decrease) the probability of disease across the predetermined threshold at which a new plan is chosen. For a variety of reasons, however, applying this principle may be difficult. Some physicians may not know the mechanics of determining whether a test is indicated. Others who are familiar with such analysis still may not employ it routinely. And, finally, some may apply it incorrectly (since, for example, the nonlinear relationship between pretest probability and posttest probability introduces error easily, unless explicit calculations are performed). TICs may be used to introduce the analysis of test indications to those not yet familiar with such an approach, and to add ease and accuracy to the work of those who are.

Test indication curves graphically demonstrate solutions that are ordinarily calculated numerically. The numerical approach proceeds as follows: values

for the pretest probability of disease and the sensitivity and specificity of the test at hand are entered into Bayesian probability equations. Solution of these equations yields values for the posttest probabilities of disease for both positive and negative test results. One then compares these new, posttest probability values with the threshold for treatment. If the posttest probability derived at a value for pretest probability (for either positive or negative results) crosses the treatment threshold, then the test is indicated *at that level of pretest probability*. If neither positive nor negative test results place the likelihood of disease on the other side of the threshold line, though, the test is clearly not indicated.

Pauker and Kassirer¹ have shown that values for the treatment threshold, coupled with the sensitivity and specificity of a test, may be used to define *testing* thresholds, and thus segregate patients into three groups. The first group includes those in whom the chance of disease is sufficiently low that neither testing nor treatment is needed. For patients in this no-test/no-treat group, the chance of disease being present *even after a positive test* still falls below the threshold for treatment. Among other patients, the no-test/treat group, the chance of disease before the test is sufficiently high that, likewise, a negative result still could not support a decision to withhold treatment. Finally, patients whose probabilities of disease lie between these extremes belong in the test/treat-if-positive group. For them, the test is indicated, for the decision to treat is correctly based

Received August 30, 1995, from the Department of Orthopedic Surgery, Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, Pennsylvania. Revision accepted for publication March 14, 1995.

Address correspondence and reprint requests to Dr. Bernstein: Department of Orthopedic Surgery, Two Silverstein Pavilion, 3400 Spruce Street, Philadelphia, PA 19104.

on the test result. TICs facilitate the identification of this important group. (As noted by Glasziou,² the test/treat-if-positive range is drawn in somewhat when the cost of the test itself is high relative to the costs and benefits of treatment.)

Determining whether a patient belongs in the test/treat-if-positive group may be difficult: testing thresholds must be determined uniquely for every treatment threshold. That is, the testing threshold is not an abstract attribute of the test itself, but rather a value for a particular patient, in a particular context. On the other hand, the functional relationship between posttest probability and pretest probability, for a given test sensitivity and specificity, is globally applicable to all patients and sets of treatment thresholds.

Test-indication curves are accordingly drawn in two parts. The first part represents the constants of the test. These "raw curves" plot posttest probabilities of disease for both positive and negative test results as a function of pretest probability. They make no reference to a particular treatment threshold, and can therefore be used universally, for all patients. The second part of a TIC, the "final curve," adds three straight lines—drawn easily by hand—at points determined by a given treatment threshold. The final curve divides the range of pretest probability into the three groups described by Pauker and Kassirer.

Creating a Raw Curve

Two separate curves plotted simultaneously on the same axes comprise the raw curve. The ordinate represents the pretest probability and the abscissa, the posttest probability. The lines of the curve are the solutions to the Bayesian equations for posttest probability, solved for the complete range of pretest probability, 0 to 100%.

These curves may be plotted employing spreadsheet software; alternatively, a simple computer program can iterate through the range of 0 to 100% and solve the equations. An example of a raw curve, for the magnetic resonance imaging (MRI) diagnosis of a meniscal tear of the knee, is shown in figure 1. Here, posttest probability as a function of pretest probability is calculated using values for sensitivity (0.93) and specificity (0.84) published in a large series in the orthopedic literature.³

Creating the Final Curve

Given a raw curve, drawing the final curve is easy. This may be done by a clinician at the bedside, using only a pencil and a straight edge, and his knowledge of the treatment threshold. The treatment threshold,

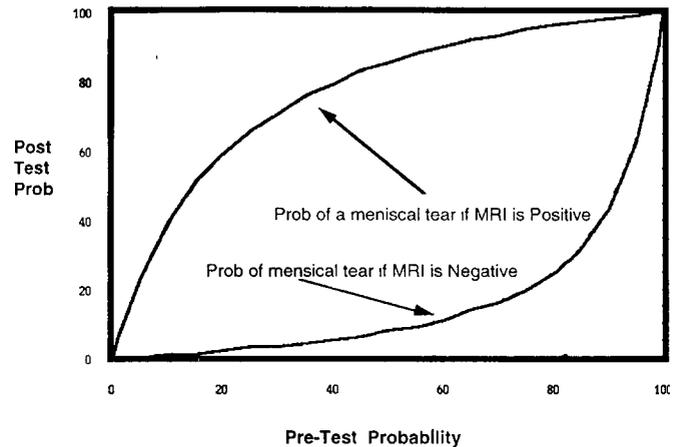


FIGURE 1. The raw curve. This is a simultaneous plot of posttest probabilities for both positive and negative test results as a function of pretest probability.

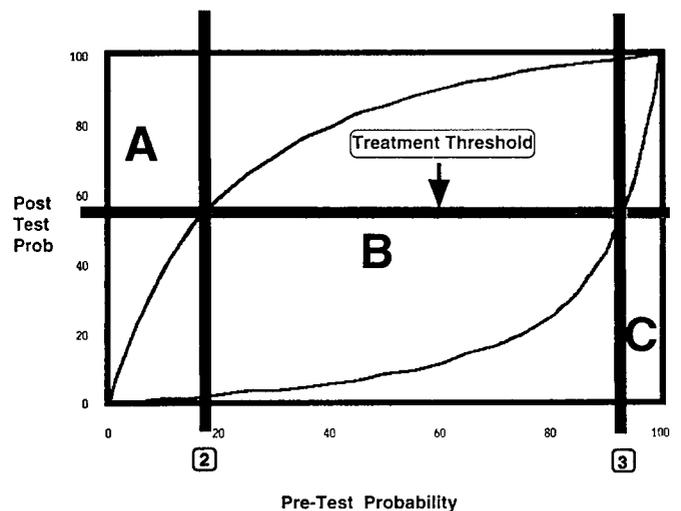


FIGURE 2. The final curve. A horizontal line (line 1) is drawn across the curve at the chosen value for the treatment threshold. The points where the drawn horizontal line intersects the raw curve are noted. Two additional lines (2 and 3) are hand drawn at these points. The range of pretest probability is now divided into three zones: A, no-test/no-treat; B, test/treat-if-positive, and C, no-test/treat. The points where the lines intersect the x-axis give the values defining the testing thresholds.

as defined by Pauker and Kassirer, is the ratio of costs of treatment to the sum of the costs and the benefits—the more costly the treatment, the more diagnostic certainty demanded. Since the threshold is calculated as a ratio, determining its value demands an assessment of only the relative magnitudes of these two values; absolute figures are not required. The treatment threshold must be established prior to the creation of the final curve. (Of course, determining this threshold is a prerequisite of any legitimate analysis of test indications, not only the creation of TICs.) In the example given, the treat-

ment threshold was 55%, the mean value found in a study of 25 active, healthy young women volunteers who were solicited from the university community. They were presented with a clinical scenario, and noted their values on a visual-analog scale. By measuring the perceived utilities for both surgical and conservative options, the treatment threshold was derived. In clinical practice, this threshold may be obtained from varied sources, such as published standards, complication rate data, or doctor-patient discussion, among others.

Two steps are needed to create the final curve from the raw curve:

Step 1. Draw a horizontal line across the raw curve at the point on the y-axis corresponding to the threshold for treatment (line 1).

Step 2. Note the points where the line drawn in step 1 intersects the raw curve, and at these two points of intersection, draw vertical lines, perpendicular to the x-axis (lines 2 and 3).

These two vertical lines now divide the raw curve into three regions. The left-most region, marked A in figure 2, is the no-test/no-treat zone. For patients for whom the probability of a meniscal tear lies in this zone, a positive MRI fails to increase the likelihood of disease above the 55% treatment threshold. Even with a positive MRI, these patients indeed should be treated as if no tear were present. (A negative test would, of course, point to no treatment as well.) Thus, the MRI examination is not indicated for them: no new clinical plan could be reasonably effected by the test. In the rightmost region, marked C in figure 2, we find the no-test/treat group. For them, a negative MRI is not enough to dissuade the surgeon from performing arthroscopy, as the post-test probability for a negative MRI is still above the treatment threshold. Only in the central group, B in figure 2, is MRI indicated. For them, the decision to operate is indeed based on the test result.

Discussion

A TIC promotes the valid use of diagnostic tests. As a visual representation of Bayesian equations, it demonstrates (rather than declares) why a test may not be indicated. It applies the work of Pauker and Kassirer without mathematics at the bedside, demanding of the end-user the addition of only three straight lines.

One may complain that TICs are meaningless should the clinician fail to establish the treatment threshold. This "weakness" may indeed be another

strength, in that it makes explicit a requirement for proper indications analysis—that of defining the treatment threshold—that might otherwise be ignored.

Creating TICs requires solving Bayesian equations, a perhaps tedious task. That requirement, however, should not undermine the use of these curves. To start, any and every application of the Pauker and Kassirer method requires calculations. Also, by separating the generally applicable and repeatable aspects of the diagnostic test (as displayed in the raw curve) from that which is unique to a given patient setting, TICs may promote greater usage of Pauker-and-Kassirer-style analysis. That is, while the fixed cost of the TIC method—creating the raw curve—may be higher than that of calculating testing thresholds for one patient, the marginal cost of applying it to additional patients is clearly substantially lower. Also, labor might be divided in a way that further encourages the use of such analysis: the creation of the raw curve might be a one-time task of the researcher who reports the sensitivity and specificity data, whereas the simpler task of drawing straight lines is reserved for the clinician.

Of course, the treatment threshold used in this example, 55%, may belong only to young, active patients. Since this threshold is defined in terms of perceived costs and benefits, should a patient not particularly value the benefits of partial meniscectomy, the treatment threshold would be high. For example, an older or more sedentary patient might not object to a period of watchful waiting, in which knees with no meniscal tear improve spontaneously. For them, a considerably higher treatment threshold is appropriate. On the other hand, the treatment threshold for a professional athlete, with a different set of cost and benefit estimates, would be even lower. For both groups, the final curve would be different than the one drawn in figure 2. This, however, is an easy modification, as the "difficult" portion, the raw curve, is universally applicable.

Finally, there are times when the precise treatment threshold is not known, yet a reasonable approximation can be made. In those cases, TICs are still helpful. By drawing lines representing the upper and lower bounds of the treatment threshold on the same raw curve, the clinician may identify the regions where the test is definitely not indicated, as well as the two "gray zones"—regions where the test perhaps might be indicated, but only if the treatment threshold were more exactly defined. In this manner, the clinician may reserve precise localization of the treatment threshold for only those patients for whom it would make a practical difference.

Test-indication curves synthesize the work of Fagen,⁴ whose nomogram initiated the move toward

the easy clinical use of Bayes' theorem; Glasziou,² who applied the Pauker-and-Kassirer analysis to the Fagen nomogram; and Daniel and Daniel,⁵ who employed the graphic representation of posttest probabilities to facilitate the comprehension of the non-linear relationship between pre- and posttest probabilities. All of these works aim to limit diagnostic testing to precisely those patients for whom it is indicated. This approach helps spare the patient the cost, burden, and risk of superfluous testing, and helps spare the physician the cost, burden, and risk of interpreting inconclusive test results.

References

1. Pauker SG, Kassirer JP. The threshold approach to clinical decision making. *N Engl J Med.* 1980; 302:1109-17.
2. Glasziou P. Threshold analysis via the Bayes nomogram [letter]. *Med Decis Making.* 1991; 11:61-2.
3. Fischer SP, Fox JM, Del Plazzo W, Friedman MJ, Snyder SJ, Ferkel RD. Accuracy of diagnoses from magnetic resonance imaging of the knee. A multi-center analysis of one thousand and fourteen patients. *J Bone Joint Surg.* 1993; 73-A:2-10.
4. Fagen TJ. Nomogram for Bayes' formula. *N Engl J Med.* 1975; 293:257.
5. Daniel BL, Daniel TM. Graphic representation of numerically calculated predictive values. *Med Decis Making.* 1993; 13:355-8.

THANKS TO REVIEWERS

MDM and the Editors thank our manuscript reviewers for *Medical Decision Making*

September 1, 1995–August 31, 1996

Daniel A. Albert, MD	Vivek Goel, MD	Harold Lehmann, MD, PhD	Mark S. Roberts, MD
David A. Asch, MD	Mary Kane Goldstein, MD	Mark Liebow, MD, MPH	Marilyn L. Rothert, PhD
Jonathan Baron, PhD	Robert M. Golub, MD	Tracy Lieu, MD	David R. Rovner, MD
William C. Baxt, MD	Philip Goodman, MD	Joseph Lipscomb, PhD	Francois Sainfort, PhD
J. Robert Beck, MD	Lee A. Green, MD, MPH	Benjamin Littenberg, MD	Susanne Salem-Schatz, ScD
George R. Bergus, MD	Robert Griffiths, MS, ScD	George Loewenstein, MD	Marilyn Schapira, MD, MPH
Marilyn Bogner, PhD	Larry Gruppen, PhD	Bryan R. Luce, PhD	Gordon Schiff, MD
Patrick Bossuyt, PhD	Robert Hamm, PhD	Jeanne Mandelblatt, MD, MPH	Henk Schmidt, PhD
Edward J. Boyko, MD, MPH	James A. Hanley, PhD	David B. Matchar, MD	Gerald Schumacher, PharmD
Robert Bryg, MD	Ron D. Hays, PhD	Dennis Mazur, MD, PhD	Carolyn Schwartz, ScD
Kenneth Buckingham, PhD	Paul Heckerling, MD	Bruce D. McCarthy, MD	Barbara Sharf, PhD
Jan Roelf Bult, PhD	Mark Helfand, MD	Donna Katzman McClish, PhD	Michael Shwartz, PhD
John Cairns, PhD	John C. Hershey, PhD	Douglas C. McCrory, MD	Marc D. Silverstein, MD
Scott B. Cantor, PhD	David Hickam, MD	Jon F. Merz, JD, PhD	Laura Simonoff, PhD
Randall D. Cebul, MD	Joergen Hilden, MD	Charles E. Metz, PhD	Wally R. Smith, MD
Robert M. Centor, MD	Margaret M. Holmes-Rovner, PhD	John Miyamoto, PhD	Eugene Somoza, MD, PhD
Gretchen Chapman, PhD	David Holtgrave, PhD	Alan J. Moskowitz, MD	Frank A. Sonnenberg, MD
Caryn Christensen, PhD	John Hornberger, MD	Essy Mozaffari, PharmD	Stephen J. Spann, MD
Brian Cohen, MD	Maria G. M. Hunink, MD, PhD	John Murphy, MD	Theodore Speroff, PhD
Christopher Colenda, MD	Thomas F. Imperiale, MD	C. David Naylor, PhD	Nils Olov Stalhammar, PhD
Alfred F. Connors, Jr., MD	Holly Jimison, PhD	Robert F. Nease, Jr., PhD	Ronald H. Stevens, PhD
Shawn Curley, PhD	Magnus Johannesson, PhD	Peter Neuman, ScD	David Stier, MD
Israel David, PhD	Robert Jones, MD, MBA	Erik Nord, PhD	Anne Stiggelbout, PhD
J. C. J. M. DeHaes, PhD	Lawrence Joseph, PhD	Bernard O'Brien, PhD	Sean D. Sullivan, PhD
Michael L. Dekay, PhD	Robert M. Kaplen, PhD	Annette O'Connor, RN, PhD	Walton Sumner, MD
Allan S. Detsky, MD	Michael Kattan, PhD	Lucila Ohno-Machado, MD	Heather Sutherland, PhD
George A. Diamond, MD	Emmett Keeler, PhD	Douglas K. Owens, MD	Tammy Tengs, ScD
Stephen Downs, MD, MS	Jeff Keisler, PhD	Christina Palmer, PhD	Danielle Timmermans, PhD
Clay Dunagan, MD	John Kellett, MD	Stephen G. Pauker, MD	George Torrance, PhD
Mark Ebell, MD	Gwendoline Kiebert, PhD	John C. Peirce, MD	Joel Tsevat, MD
Louis Eeckhoudt, PhD	Job Kievit, MD, PhD	James Pellissier, PhD	Peter A. Ubel, MD
Linda Emanuel, MD	Bruce Kinosian, MD	Will Penny, PhD	Jerome Van Ruiswyk, MD, MS
Jeff Etchason, MD	Sara J. Knight, PhD	Peter Politser, MD	Willem Vandaal, PhD
Ann Maxwell Eward, MD	J. André Knottnerus, MD, PhD	Roy M. Poses, MD	Lia C. G. Verhoef, MD
Ellen Fox, MD	Michael Koller, PhD	Gad Rabinowitz, PhD	John Vessey, PhD
Kenneth Freedberg, MD	Murray Krahn, MD	David F. Ransohoff, MD	Douglas P. Wagner, PhD
Theodore G. Ganiats, MD	Karen Kuntz, ScD	Michael M. Ravitch, PhD	Peter Wakker, PhD
Joseph Gardiner, PhD	Curtis P. Langlotz, PhD	Donald Redelmeier, MD	Elke Weber, PhD
Constantine Gatsonis, MD	William F. Lawrence, MD, MS	Dennis A. Revicki, PhD	Richard Willke, PhD
David Gift, MS, SM		Brian E. Rittenhouse, PhD	Fredric M. Wolf, MD
Paul Glasziou, PhD			John B. Wong, MD
Henry Glick, MD			John R. Woods, PhD