

# Visual Presentation of Statistical Concepts in Diagnostic Testing: The 2 × 2 Diagram

Kevin M. Johnson<sup>1</sup>  
Benjamin K. Johnson<sup>2</sup>

**OBJECTIVE.** The purpose of this article is to present a visual conceptual framework for important statistical concepts in radiology, and to provide an online application to facilitate this visualization.

**CONCLUSION.** Statistical measures such as sensitivity, specificity, and predictive values are ubiquitous in medical literature, yet thinking fluidly about these concepts is not always easy. The 2 × 2 diagram is a helpful guide.

**S**tatistical descriptions find application far afield of the professional precincts in which they are most fully understood. They are generated by experts but must be understood by amateurs. In medicine, statistics serves as both a test of meaningfulness and a form of persuasion. But even an elegant analysis is of little use if not understood by its nonstatistician readers.

Results of diagnostic studies are commonly described with 2 × 2 contingency (or truth) tables. However, these tables can behave in complicated ways. In a previous article [1], we introduced a graphic interpretation of these tables called a 2 × 2 diagram, which allows straightforward mental manipulation of the data. Standard statistics—including sensitivity, specificity, and predictive values—can be seen at a glance. In the current article, we extend the diagram to illustrate likelihood ratios, odds ratio, and expected values, the Pearson chi-square statistic, Fisher exact test, and ROC curves. We also provide a link to a web application, with which readers can explore and generate diagrams for their own teaching and research purposes.

## The 2 × 2 Table

Many problems in diagnosis are treated as a two-valued logic problem: A disorder is considered to be either present or absent, and a test for the disorder is considered to have either positive or negative results. This is an oversimplification, but nevertheless finds widespread use. Assuming that the truth is known by some independent means,

a given test result falls into one of four possible categories: false-positive (FP)—test result is positive, disorder is truly absent; true-positive (TP)—test result is positive, disorder is truly present; true-negative (TN)—test result is negative, disorder is truly absent; and false-negative (FN)—test result is negative, disorder is truly present. These results can be presented as in Figure 1, variously termed a contingency table, truth table, 2 × 2 table, or decision matrix. Familiar statistics are defined as proportions of the four basic results; for example, sensitivity is defined as  $TP / (TP + FN)$ , and specificity as  $TN / (TN + FP)$ .

In what sense is this table complicated [2, 3]? What does a particular distribution of cell values imply for the usefulness of the test prospectively? When is a difference between cells significant, that is, unlikely to have arisen by chance alone? To answer these and other questions, a number of descriptive methods have emerged. A good description should give an intuitive grasp of the connectedness of the cells—a property that the classic 2 × 2 table lacks. The 2 × 2 diagram was developed to fill this need.

## The 2 × 2 Diagram

The principal components of the diagram are a box to represent the subjects; a four-direction coordinate system to represent the true-positive, false-positive, false-negative, and true-negative categories; and optionally, a trajectory to represent how the box moves with changes in the threshold for abnormality. The box width is determined by the num-

**Keywords:** contingency table, diagnostic testing, predictive values, sensitivity, specificity, statistics

DOI:10.2214/AJR.13.11954

Received September 18, 2013; accepted after revision November 17, 2013.

<sup>1</sup>Department of Diagnostic Radiology, Yale University School of Medicine, 55 Lock St, New Haven, CT 06511. Address correspondence to K. M. Johnson (kevin.johnson@yale.edu).

<sup>2</sup>Yale Center for Analytical Sciences, New Haven, CT.

## WEB

This is a web exclusive article.

AJR 2014; 203:W14–W20

0361–803X/14/2031–W14

© American Roentgen Ray Society

## 2 × 2 Diagram

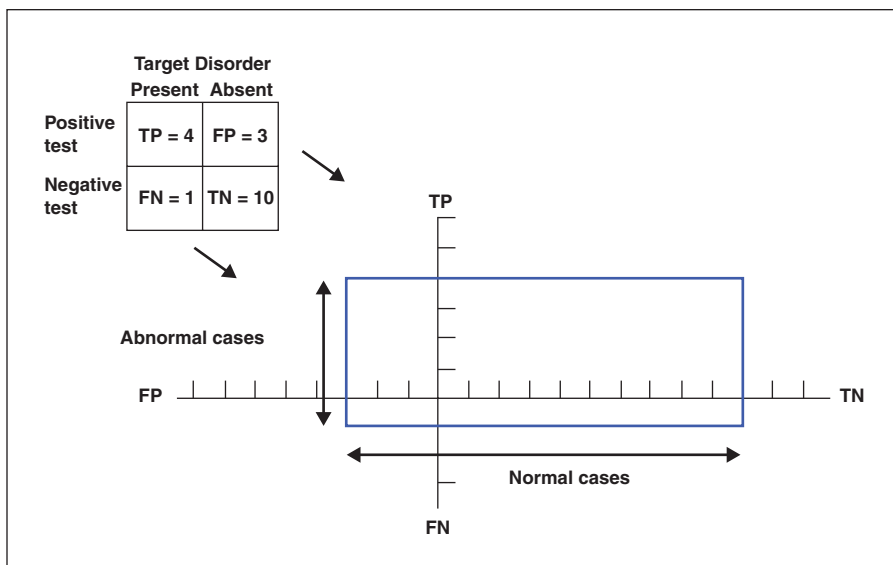
ber of healthy subjects, and the box height is determined by the number of subjects with abnormalities (Fig. 1). The prevalence of the disorder in the studied population is thus the ratio of the height of the box to the sum of its height and width. For example, a square box denotes a prevalence of 50%, a short, wide box denotes a low prevalence, and a tall, narrow box denotes a high prevalence.

Each hemiaxis of the coordinate system represents one of the four cells of the contingency table. Subjects with abnormalities lie along the vertical axis: the upper hemiaxis represents the true-positive results, and the lower hemiaxis represents the false-negative results. Healthy subjects are along the horizontal axis: the left hemiaxis represents the false-positive results, and the right hemiaxis represents the true-negative results.

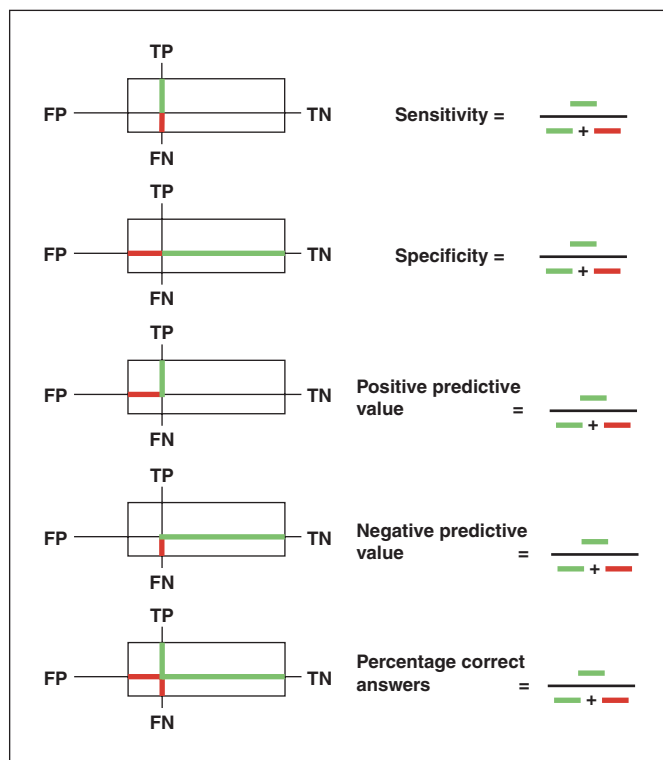
The position of the box is a function of the behavior of the diagnostic test (Fig. 2). If sensitivity is high, most of the box lies above the horizontal axis; if specificity is high, most of the box lies to the right of the vertical axis. Thus the box corresponding to a perfectly accurate test would lie entirely in the upper right quadrant, indicating 100% sensitivity and 100% specificity. Positive predictive value (PPV) is given by the proportions of the left upper quadrant of the box, where only positive results reside. Because PPV is defined as the ratio of true-positive results to total positive results, a tall, narrow quadrant denotes a high PPV. Likewise, the negative predictive value (NPV) is defined as the ratio of true-negative results to total negative results, and is thus represented by the proportions of the right lower quadrant of the box. A short, wide quadrant denotes a high NPV.

### Attributes of a Good Test

A good test places much of the area of the box into the right upper quadrant, thus reflecting high sensitivity and specificity. Imagine the box moving around on the coordinate system. As more of the box moves into the ideal right upper quadrant, the sensitivity is increasing (box moving upward), the specificity is increasing (box moving rightward), the PPV is increasing (left upper quadrant becoming taller and narrower and eventually a vertical line), and the NPV is also increasing (right lower quadrant becoming wider and shorter and eventually a horizontal line). The product of the sensitivity and specificity is the proportion of the total area of the box that lies in the right upper



**Fig. 1**—Two by two (2 × 2) table and diagram. Construction of table begins with determination of actual presence or absence of target disorder according to some reference standard. Test is performed, and results are classified as either positive or negative based on specific criteria. Results then fall into four categories: true-positive (TP), false-positive (FP), false-negative (FN), and true-negative (TN). Two by two diagram is graphic transformation of table. Box represents subjects with abnormalities and healthy subjects; coordinate system represents each of four result categories.



**Fig. 2**—Common terms in 2 × 2 diagram. In each case, length of green line is divided by sum of lengths of green and red lines. Imagine box as moving around on coordinates; various proportions change in concert, not independently. TP = true-positive, FP = false-positive, FN = false-negative, TN = true-negative.

quadrant. Because sensitivity and specificity are often emphasized in clinical reports, we have found this product useful, along with an estimate of the prevalence, to roughly visualize the diagram while reading the literature.

Sensitivity and specificity are not independent. As sensitivity increases and the box moves upward, it almost always also moves leftward, decreasing the specificity. This path can be defined by a curved line in the

left upper quadrant, along which the left upper corner of the box moves. We call this the test trajectory (yellow line, Fig. 3).

If the test has no discriminatory value, knowing the test results does not change the probability of having the disease; that is, the posttest probability equals the pretest probability. In truth diagram terms, this means the proportions of the left upper quadrant of the box (the posttest probability) are identical to the proportions of the box as a whole (the pretest probability). This criterion is met by every position of the box when the test trajectory follows the diagonal line (red line, Fig. 3). In this case we give the test trajectory a special name,

the chance trajectory, for a test with no discriminatory value.

**Likelihood Ratios and the Odds Ratio**

In addition to statistics such as sensitivity and specificity, other measures have been developed to try to summarize how well a test performs. Two of these are likelihood ratios and odds ratio.

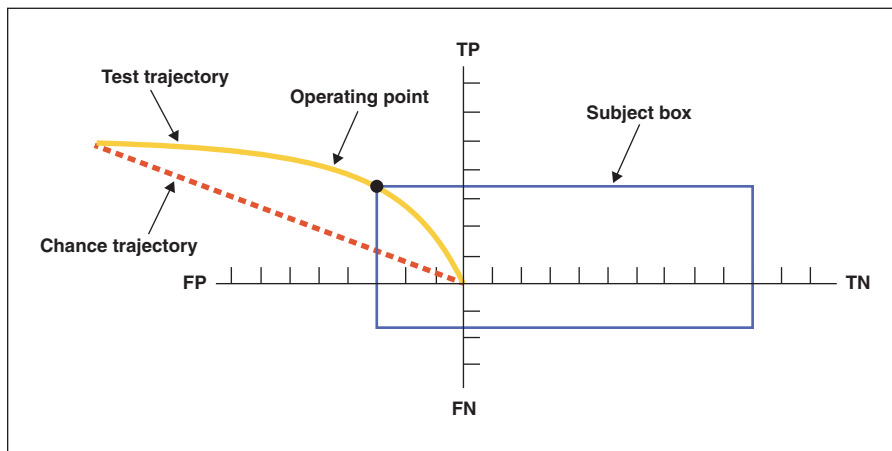
We can define three relevant odds for a diagnostic test: the odds of a disease being present given a positive test result, the odds of a disease being present given a negative test result, and the odds of disease being present before testing. In the context of the 2 x 2 diagram, odds are represented by

slopes (Fig. 4). The three possible ratios of these odds can be used to summarize the performance of the test: positive likelihood ratio (+LR) = odds given positive test result / prior odds (i.e., prevalence); negative likelihood ratio (-LR) = odds given a negative test / prior odds (i.e., prevalence); odds ratio (OR) = odds given a positive test result / odds given a negative test result.

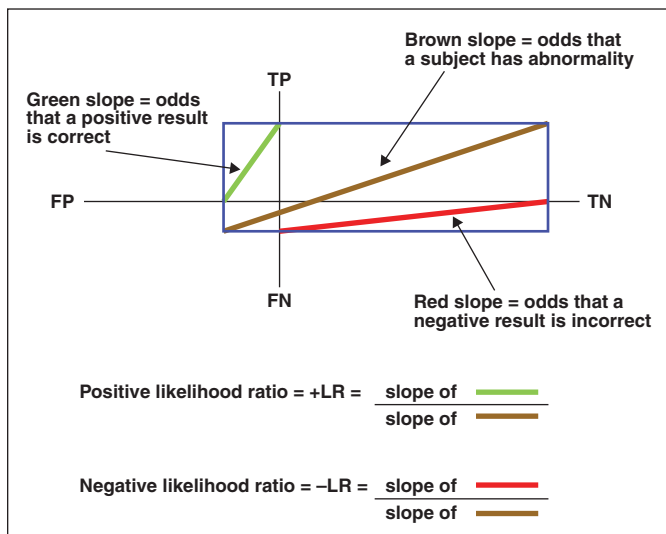
The slope in the left upper quadrant indicates the odds of disease given a positive test result; the slope in the lower right quadrant indicates the odds of disease given a negative test result; and the slope of the entire box gives the odds of disease before testing (i.e., the prevalence). Thus, we see that the positive likelihood ratio is high when the left upper quadrant is tall and narrow and that the negative likelihood ratio is low when the right lower quadrant is short and wide. This is consistent with the idea that test performance improves as the box moves into the right upper quadrant.

The likelihood ratios multiply the odds. For example, if the pretest odds are 1:10 and the positive likelihood ratio is 10, the posttest odds of having the disorder given a positive test result are 1:1. If the pretest odds are 1:10 and the negative likelihood ratio is 0.1, the posttest odds of having the disorder given a negative test result are 1:100.

The odds ratio is the ratio of the odds that a disease is present given a positive test result to the odds that disease is present given a negative test result. Thus odds ratio = +LR / -LR. However, a high odds ratio can be seen even though the test has poor specificity and may not be clinically useful; it does not necessarily mean that most of the box is in the right upper quadrant (Fig. 5). Any given odds ratio defines more than one possible position for the box and so does not describe a unique 2 x 2 diagram even if the box size and proportions are known. Odds ratios should be used with caution for these reasons. This is an example in which the diagram reveals a limitation of a summary statistic.



**Fig. 3**—Test and chance trajectories. Test trajectory (yellow) is curve along which left upper corner of blue box is constrained to move. Point of intersection with trajectory is called operating point. Box is attached at operating point and slides along trajectory as cut point, or test threshold of abnormality, changes. Each side of box can reach but never cross axis parallel to it, so origin of axes is always within or on edge of box. ROC curve can be constructed from trajectory. Chance trajectory (red) is trajectory along which posttest probability of disease is numerically equal to pretest probability, that is, trajectory along which test does not alter probability of disease and is therefore worthless. TP = true-positive, FP = false-positive, FN = false-negative, TN = true-negative.



**Fig. 4**—Likelihood ratio and odds ratio. Slope values represent odds as shown. Ratios of these odds can be used as summary statistics. For example, positive likelihood ratio (+LR) of 5 means that slope of green line is five times that of brown line. If test result is positive, odds of patient having disease have increased by factor of 5 over pretest odds. TP = true-positive, FP = false-positive, FN = false-negative, TN = true-negative.

**Expected Values**

The underlying concept of expected values must be described before the way in which the 2 x 2 diagram illustrates the Pearson chi-square test can be delineated. “Expected values” refers to the set of values for the 2 x 2 table that we would expect to see if the results of the test were completely independent from the presence of disease. In estimating expected values, we calculate the

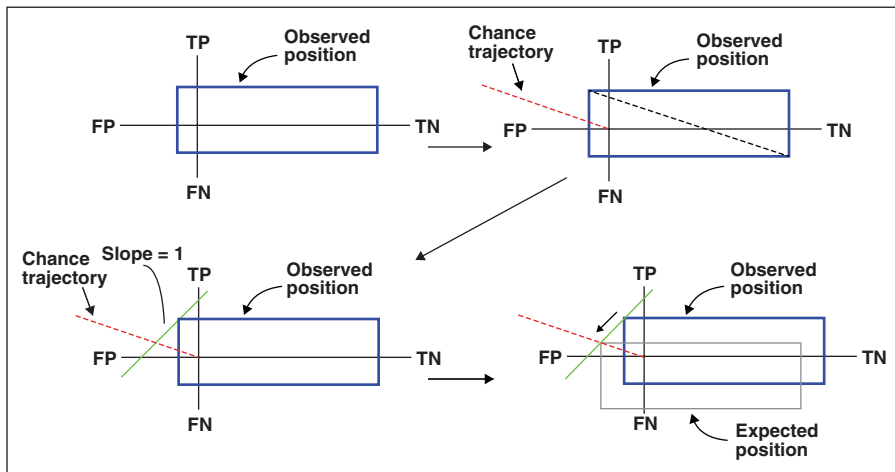
## 2 × 2 Diagram

arrangement of cells in which the same proportions of patients with disease and healthy persons receive a positive test result while maintaining the same ratios of healthy persons to patients with disease and negative to positive test results (i.e., the 2 × 2 table marginal totals remain fixed). On the diagram this translates to the following: The size and proportions of the box remain unchanged, and the length of the periphery of the upper left quadrant of the box remains unchanged, that is, the sum TP + FP stays constant (as does the sum TN + FN).

Thus to show expected values, the box (solid blue, Fig. 5) moves from the observed position toward the chance trajectory (red line, Fig. 5) without a change in the total number of results in the left upper quadrant (TP + FP). This means that the left upper corner of the box must move along the green line because that is the only path possible if TP + FP is kept constant. The expected position of the box is then the position where the left upper corner sits at the intersection of this green line and the chance trajectory. We call the intersection of the box with a trajectory the operating point. Any operating point along the chance trajectory can occur by chance, but only this point also satisfies the fixed-margin condition.

### Pearson Chi-Square Test

Once the expected position of the box is calculated, we can quantify the distance between the actual box position and the expected position. A position off the chance trajectory position is possibly the consequence of an actual power of the test to discriminate healthy subjects from subjects with an abnormality, and possibly the consequence of chance. Thus, specifically, we seek a meth-



**Fig. 5**— Finding expected position of box. Box position expected by chance alone can be computed from actually observed position: Box moves from observed position (solid blue) toward chance trajectory (red) without changing total number of results in left upper quadrant, that is, total number of positive results (TP + FP). This means that left upper corner of box must move along green line of slope = 1. Expected position of box is with its left upper corner at intersection of this green line and chance trajectory. Any operating point along chance trajectory could occur by chance, but only this point also satisfies additional (by convention) requirement that total number of positive results (and total number of negative results) remains unchanged. TP = true-positive, FP = false-positive, FN = false-negative, TN = true-negative.

od for determining the probability that the observed separation is the result of chance alone. The Pearson chi-square statistic is a measure of this probability. The null hypothesis is that the observed box is merely the expected box, which has by chance wandered away from its mean position. This wandering is presumed to (approximately) obey a normal probability distribution along each of the four axis directions. The farther apart the two boxes, the less likely it is that the separation has been caused by random variation.

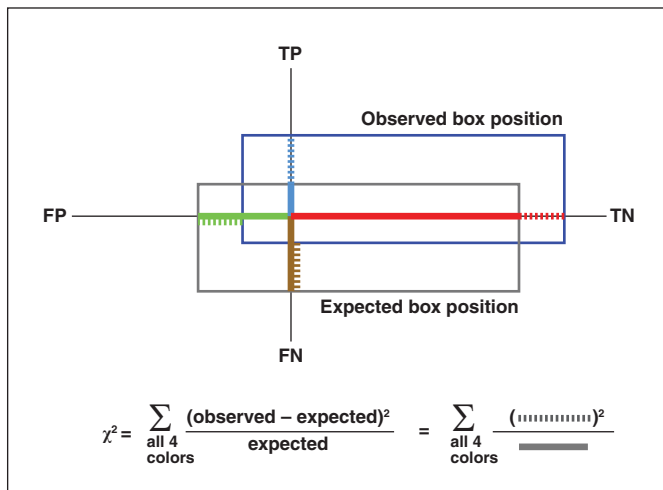
The Pearson chi-square statistic is calculated from the 2 × 2 diagram as follows: For each of the four directions, the displacement between the expected and observed boxes is squared and normalized by the expected

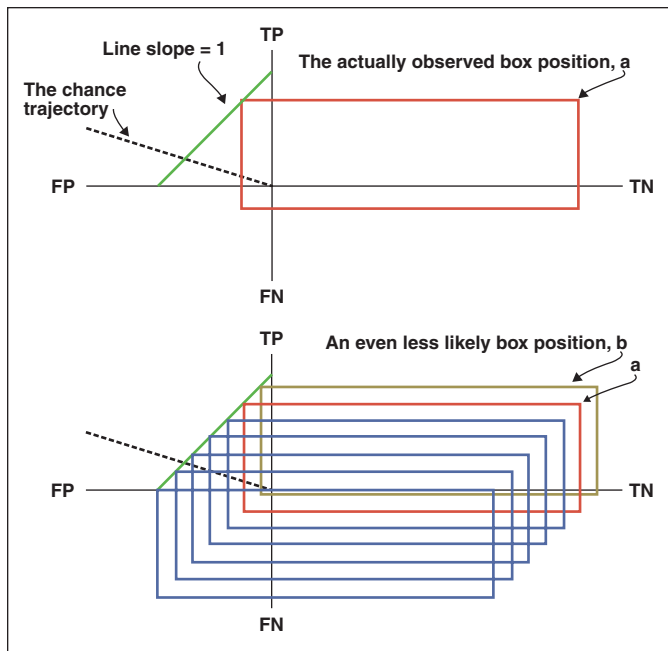
value associated with that direction (Fig. 6). The resulting four values are summed, yielding the Pearson chi-square statistic. Because this value is the sum of squares of normally distributed values, it follows the chi-square distribution. The statistic is interpreted by use of a lookup table for the chi-square distribution with one degree of freedom. The *p* value obtained from the table is the probability that the observed box would be found at least this far away from its expected position if chance alone were operating.

### Fisher Exact Test

The 2 × 2 diagram can also be used to visualize the Fisher exact test, an alternative to the Pearson chi-square test that is particularly

**Fig. 6**— Pearson chi-square. First, position of subject box is observed on basis of actual results of diagnostic test. Next, box position expected by chance alone is determined. Then for each hemiaxis, distance between two boxes is squared and normalized by expected value for that hemiaxis. Resulting four terms are summed to give chi-square value; reference to table of chi-square distribution with one degree of freedom returns *p* value. If chi-square exceeds 3.84, there is only 5% chance that observed box position has occurred by chance alone. TP = true-positive, FP = false-positive, FN = false-negative, TN = true-negative.





**Fig. 7**— Fisher exact test. First, box is constrained to move along green line by assumption that marginal totals of contingency table remain constant. All possible box positions are considered; their axis positions can take integer values only. Second, Fisher exact test calculates probability of each position of box occurring by chance using hypergeometric distribution. This probability declines with distance from chance trajectory. Third, probability for original (observed) position is added to probabilities of even more unlikely positions (*brown box*). If this sum of probabilities is small, for example, < 0.05, null hypothesis that results are from chance alone is rejected. TP = true-positive, FP = false-positive, FN = false-negative, TN = true-negative.

useful when one of the table cells has a small number of observations ( $n < 5$  is a common cutoff). In this test, the observed  $2 \times 2$  table is considered a member of all  $2 \times 2$  tables with the same marginal totals. Fisher showed that the hypergeometric distribution can be used to calculate the probability of each box position under the null hypothesis that the diagnostic test has no discriminatory power. Specifically, the probability of each position is given by

$$p = \frac{(r1! \times r2!Lrc!) \times (c1! \times c2!Lcr!)}{(x11! \times x12!Lxrc! \times N!)}$$

where  $ri$  is the sum of the  $i$ th row of the table,  $cj$  is the sum of the  $j$ th column of the table, and  $N$  is the total number of observations. If the probability of all tables at least as extreme as the observed table is small (i.e.,  $p < 0.05$ ) the null hypothesis is rejected.

In the  $2 \times 2$  diagram context, this family of tables is represented by all boxes of the

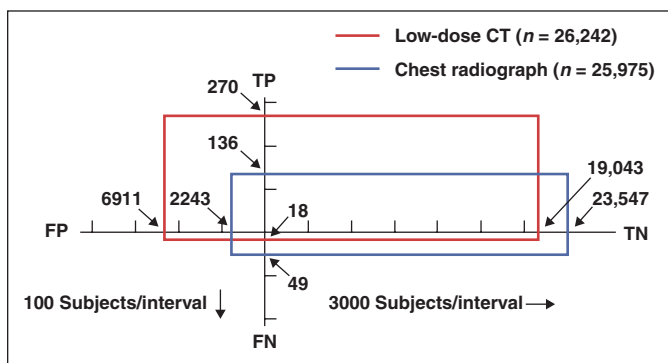
same size with upper left corners that lie on the green line (Fig. 7). The boxes that lie farther toward the upper-right quadrant represent tests with better discriminatory power and are farther from the chance trajectory and thus less likely to occur by chance. The Fisher exact test consists of adding the probability of the observed position to the probabilities of all positions that lie farther up and to the right of the observed position and determining whether this probability is sufficiently small.

**Test Trajectory Versus Receiver Operating Curve**

Many, though not all, diagnostic tests entail some kind of continuous measure for the condition under consideration. Most often we want to think of a condition or disease as being present or not present, which is a dichotomous choice, so we have to assign a cutoff value of some kind to separate the sample results into two parts. But the choice

of cutoff value is often somewhat arbitrary, and it is of interest to be able to describe the test behavior over a range of cutoff values, not just a single one. A single cutoff value gives a single  $2 \times 2$  table. A range of cutoff values requires multiple tables, each resulting in a different position of the box on the  $2 \times 2$  diagram. Each position of the box has a different point for the left upper corner, and all of these possible points taken together define a line we call the test trajectory (Fig. 3). It is the path taken by the box as the cutoff criterion is varied. This is seldom a straight line. Sensitivity and specificity are not independent: As the threshold for abnormality is changed, the proportions of correctly classified healthy subjects and subjects with abnormalities change in concert.

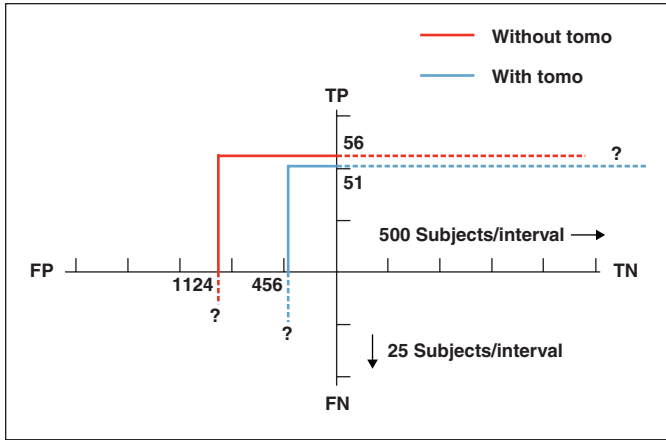
ROC curves and trajectories are constructed in a similar way, except that ROC curves ignore prevalence. Sensitivity and  $1 -$  specificity are calculated with each value of the



**Fig. 8**— National Lung Cancer Screening Trial. Low-dose CT was compared with chest radiography for detection of lung cancer. Results are for first year of screening [5]. Y-axis is magnified 30x relative to x-axis. CT has higher sensitivity but at cost of lower specificity, as shown by its position shifted upward and to left relative to radiography. Even though two subject groups were carefully matched for risk, number of abnormal findings (vertical dimensions) in radiography group was substantially less than in CT group. Because detection of abnormal findings was in part dependent on corresponding technique under study, less sensitive technique (radiography) likely missed number of subclinical cancers, which if detected would have resulted in two boxes having similar proportions. Subjects classified as having true-negative (TN) results would be reclassified as having false-negative (FN) results. Box proportions would become taller and narrower, but because left upper corner of revised blue box would not move, more of box would be below x-axis than in original analysis. True sensitivity of radiography would be worse than shown by original diagram. Specificity would be better. TP = true-positive, FP = false-positive.



## 2 × 2 Diagram



**Fig. 9**—Digital mammography with and without tomosynthesis (tomo). Y-axis is magnified 20× relative to x-axis. Dashed lines indicate that rest of box dimensions and corner positions are unknown. Only left upper quadrant can be drawn from published data [7]. What we know: Absolute positive predictive value has increased from 4.7% to 10.1%, sensitivity is little changed, specificity has increased, positive likelihood ratio has increased, and odds ratio has increased. What we do not know: actual sensitivity, specificity, odds ratio, positive likelihood ratio, or their percentage change. TP = true-positive, FP = false-positive, FN = false-negative, TN = true-negative.

cutoff criterion and then graphed against one another. The test trajectory becomes a graph of the true-positive results versus the false-positive results. One can convert the true-positives into sensitivity simply by knowing the total number of subjects with abnormalities and likewise convert the false-positives into 1 – specificity by knowing the number of subjects with abnormalities. Thus to convert a test trajectory into an ROC curve simply requires knowing the numbers of healthy subjects and subjects with abnormalities. If we

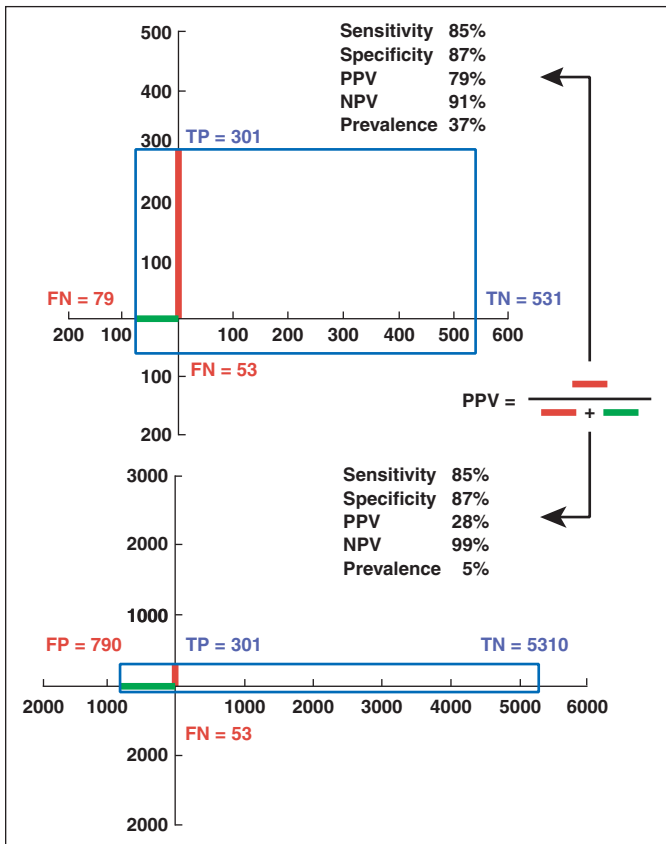
have the test trajectory, we can always construct the corresponding unique ROC curve. If instead we know only the ROC curve, we need to be told the numbers of healthy subjects and subjects with abnormalities to draw the corresponding test trajectory. In this sense the complete 2 × 2 diagram with a test trajectory contains more information than does the ROC curve.

It has been said [4] that ROC curves measure test performance independently of the study population, whereas predictive values

(and therefore our diagram) require knowledge of the particular population studied. In this view, the insensitivity of ROC curves to prevalence is a virtue. However, this means ignoring the predictive values, and that can lead to serious practical limitations in comparing different diagnostic methods unless the prevalence of the disorder is the same for the two tests. The tendency of ROC curves to obscure the influence of prevalence can be a liability.

### Example I: Lung Cancer Screening

The National Lung Cancer Screening Trial [5, 6] compared groups of subjects screened for lung cancer with either low-dose CT or single-view chest radiographs. The 2 × 2 diagrams show that CT has a higher sensitivity for tumor but a lower specificity than radiography (Fig. 8). Even though the two test groups were evenly matched in multiple risk factors for lung cancer, the proportion of subjects with abnormalities in the radiography group was substantially lower than in the CT group, as shown by the difference in proportions of the subject boxes. This raises the question whether a substantial number of cancer cases were not accounted for in the radiography box, lying below the x-axis, and therefore that the true sensitivity of radiography is substantially less than the observed values. Subclinical cancers were more likely to come to light in subjects undergoing follow-up with a highly sensitive technique than with a less sensitive one. Thus a certain self-referential bias may affect the accuracy of the sensitivity estimates. Because sensitivity is linked to the rest of the diagram, estimates of odds ratio, likelihood ratios, expected values, Pearson chi-square statistic, and so on, are also affected.



**Fig. 10**—Prostate-specific antigen. Top diagram represents case-control study comparing results for men 60–69 years old with known prostate cancer (vertical axis, red) with results for control subjects without prostate cancer (horizontal axis, green) by use of prostate-specific antigen cutoff of < 4 ng/mL [8]. Bottom diagram is derived from same data by multiplying number of healthy subjects (horizontal axis) by 10 to simulate lower prevalence of disease likely in broader screening population. Relative proportion of true-negative (TN) to false-positive (FP) results is preserved (i.e., specificity is unchanged). Positive predictive value (PPV) decreases substantially as prevalence decreases. TP = true-positive, FN = false-negative, NPV = negative predictive value.

### Example 2: Digital Mammography With or Without Tomosynthesis

Tomosynthesis is a technology for increasing the accuracy of digital mammography. Figure 9 shows the partial  $2 \times 2$  diagram that can be constructed with results published by Rose et al. [7]. Because only subjects with positive results went on to further workup, including biopsy, only the left upper quadrant of the underlying, but unknown, full box can be drawn. Conclusions can be reached about some aspects of the effect of tomosynthesis on test performance, but not about other aspects.

### Example 3: Prostate-Specific Antigen

Case-control studies are a common experimental design in which a group of patients with a disorder is matched to a group of subjects who do not have the disorder but are similar in other ways, such as age and sex. Such studies are a frequent early step in evaluation of new diagnostic tests. However, good sensitivity and specificity in a case-control study can translate to poor PPV once the diagnostic test is applied to a broader screening population. Consider the use of prostate-specific antigen level for the detection of prostate cancer. In one study [8], the PPV was 79% for men 60–69 years old (cutoff prostate-specific antigen level  $< 4$  ng/mL [Fig. 10, top]). The prevalence had been set artificially, as in all case-control studies, in this instance by including just under two control subjects for each cancer patient, giving a prevalence of 37%.

What happens when the test is applied to a general population? The natural prevalence of clinically significant cancer in this larger population is likely to be lower than the arti-

cial prevalence in the case-control population. This can be visualized as addition of healthy subjects to the horizontal axis; the subject box becomes longer and proportionately flatter (Fig. 10, bottom). The proportions of the left upper quadrant box change from tall and narrow to short and wide. The false-positive results increase, driving the PPV from 79% to 28%. The high NPV of 99% is mostly a result of the low prevalence; the test itself adds little predictive value. This deterioration of predictive values is strictly a consequence of the decrease in prevalence; the sensitivity and specificity remain unchanged. It is worth noting that an analysis using ROC curves would not illustrate this point.

### Other Uses of the Diagram

By assigning a cost (or benefit) to each hemiaxis, one can use the diagram to think about how costs and benefits push and pull the box to an optimal position. The sides of the box can be rescaled by the cost per case in each hemiaxis, and the various statistics can be recalculated in that light. Verification bias, in which only certain test results trigger the use of the reference standard for verification of the suspected diagnosis, can be illustrated [1]. With modification, the diagram can be extended beyond diagnostic tests to other situations in which  $2 \times 2$  tables are used, such as drug trials.

### Conclusion

The  $2 \times 2$  diagram can help nonstatisticians think fluidly about the interactions of sensitivity, specificity, predictive values, and prevalence. Two tests can be compared by placing two different boxes on the same coordinate

system. Whether the test results are statistically significant can be determined using the Pearson chi-square statistic, built on the concept of expected values, both illustrated by the diagram. The trajectories of the tests are analogous to ROC but preserve the prevalence information. Readers may wish to diagram their own examples; an applet can be found at <http://ycas.yale.edu/2by2.aspx>. We invite readers to use the method freely, and even expand and improve on it. The source code (in the high-level language R) will be provided on request.

### References

1. Johnson KM. The two by two diagram: a graphical truth table. *J Clin Epidemiol* 1999; 52:1073–1082
2. Fleiss JL. *Statistical methods for rates and proportions*, 2nd ed. New York: Wiley, 1981
3. Swets JA, Pickett RM. *Evaluation of diagnostic systems: methods from signal detection theory*. New York: Academic Press, 1982
4. Zweig MH, Campbell G. Receiver operating characteristic plots: a fundamental evaluation tool in clinical medicine. *Clin Chem* 1993; 39:561–577
5. National Lung Screening Trial Research Team, Church TR, Black WC, Aberle DR, et al. Results of initial low-dose computed tomographic screening for lung cancer. *N Engl J Med* 2013; 368:1980–1991
6. Aberle DR, DeMello S, Berg CD, et al. Results of the two incidence screenings in the National Lung Screening Trial. *N Engl J Med* 2013; 369:920–931
7. Rose SL, Tidwell AL, Bujnoch LJ, et al. Implementation of breast tomosynthesis in a routine screening practice: an observational study. *AJR* 2013; 200:1401–1408
8. Jacobsen SJ, Bergstralh EJ, Guess HA, et al. Predictive properties of serum prostate-specific antigen testing in a community-based setting. *Arch Intern Med* 1996; 156:2462–2468