🔓 Accessible   Published by De Gruyter August 18, 2017

Using Bayes' rule in diagnostic testing: a graphical explanation

Kevin M. Johnson

Cite this   Citations   8

Abstract

In the context of diagnostic testing, concepts such as sensitivity, specificity, predictive values, likelihood ratios and more are all interconnected, but precisely how can be confusing to the nonstatistician. This paper presents a graphical explanation. Bayes' rule, or theorem, ties several of these concepts together and should have a role in careful diagnostic thinking. It too can be understood in terms of the "two-by-two diagram" presented here.

Keywords: Bayes' theorem; diagnostic tests; diagrams; likelihood ratio; odds; posttest probability; predictive value; pretest probability; sensitivity; specificity

Introduction

A fundamental question in medicine is whether a given test will be useful to address a given clinical problem. There is often more than one choice among tests, which vary in accuracy, expense and safety. Choosing among them is a major issue not just for daily clinical care but for the health care system as a whole. Certainly, it is a daily concern in radiology.

Few tests are perfectly accurate. A good test will not often miss a diagnosis and will also not often sound a false alarm. Put a different way, a good test is usually thought of as having high sensitivity and high specificity, respectively.

Sensitivity and specificity are intimately connected: when one is increased, the other tends to decrease. In fact, some tests excel at sensitivity but have poor specificity, and others have poor sensitivity but high specificity. For a test to be useful, it is not necessary for both sensitivity and specificity to be high. It is however important for the user to understand where the strength of the test lies.

A second key concept is that published sensitivity and specificity values do not, by themselves, describe how well a test will work in the real world. This is because they are determined by already knowing the answer. They answer the question: Given that we already know the truth, how accurate is the test result? However, our question is instead: Given that we know the test result, what is the truth? Put another way, we are interested in knowing the positive and negative predictive values. Sensitivity and specify are only tools to help us get to those values.

Bayes' theorem (or rule) is the mathematical relationship that ties sensitivity and specificity to predictive values. Likelihood ratios are a convenient tool to convert the former to the latter. These concepts can be illustrated graphically, as will be discussed here.

Sensitivity and specificity are not enough

Some test results are very useful when they are positive, but essentially useless when they are negative. Others can be very useful when they are negative but not when they are positive. This two-dimensional understanding of results often gets overlooked. The simple, and incorrect, tendency is to look for a positive result, and if it is not there, to go on to the next test to look for a positive result.

In fact, definite harm can result from failing to understand this point. A D dimer is a very sensitive (~99%) test for deep venous thrombosis but nonspecific (~40%) [1]. Consider a young patient with no risk factors for DVT or pulmonary embolus who presents with dyspnea. Her pretest probability of embolus is very low, less than 2%, and for that reason, a positive D dimer would be subject to a high false-positive rate. Suppose her D dimer is done nonetheless and is positive. This increases her probability of embolus to about 3.4%, still quite low. In view of a positive test however, the clinician may feel obliged to order a CT angiogram. In one such case, the images were interpreted as showing emboli. The patient was put on an anticoagulant. Later

review of the CT images disputed the original reading. In the meantime, the patient had an intracerebral bleed. Even without this terrible outcome, the expense and radiation exposure from a CTA was unnecessary. Alternatives such as close follow-up might have been considered instead.

Therefore, a test with a low specificity and a high sensitivity might be useful if it is negative, but seldom if it is positive. This is counterintuitive. Is not the virtue of this test its high sensitivity? So if it is positive, doesn't that mean the sensitivity is "working" in some sense? It does not.

The value of a high sensitivity test is largely when the result is negative (to rule out the diagnosis). Likewise, the value of a high specificity test is largely when the result is positive (to rule in the diagnosis).

However, this is an incomplete explanation. A full explanation must hold in mind both sensitivity and specificity at the same time. It also must allow us to arrive at the predictive values. How can we do this without confusing ourselves? Isn't something missing conceptually?

The missing concept can be conveyed mathematically by Bayes' theorem (or rule). It is a mathematical relationship that takes the sensitivity and specificity, combines them with the pretest probability and outputs the positive and negative predictive values. The mathematical form is given later. However, a visual form is much easier for this radiologist author to comprehend, and I offer it to my readers as a solid way to conceptualize the topic.

Sensitivity and specificity so pervade the medical literature that you can be forgiven for assuming that they are a complete description of the accuracy of a test. Predictive values are emphasized far less often.

An important caveat is that any given pair of sensitivity and specificity numbers depend on two important factors that do not enter into the Bayes rule itself. One is that these values almost always are calculated by designating some cut point such as a numerical value in the case of a laboratory test, a velocity or velocity ratio in Doppler ultrasound or a perceptual "threshold" in the case of images. The latter is influenced by the radiologist's expectation that disease is present, or in other words, once again by the pretest probability. This in turn is strongly influenced by the history available to the radiologist.

A second important caveat is that sensitivity and specificity have been established in some test population, the composition of which might not be comparable to that of the population to which you wish to apply the test. If the research was done only after eliminating candidates with certain risk factors, or only after some other preliminary tests have been performed, and your patients have not been similarly selected, the sensitivity and specificity values may be invalid for your purposes. This is the so-called "patient spectrum" problem [2, 3].

Keep in mind that when tests are used sequentially, their characteristics interact.

For example, if the first test has low sensitivity, many abnormal cases will be missed at this point and fall out of further consideration. The false negatives do not go on to the next test. If the intention is to screen for disease, this would be a very undesirable quality. The implication is that many first line tests have to be sensitive but nonspecific, e.g. D dimer. *That means that if the pretest probability is low, there will be many false-positive results.*

A visual explanation

To explain these concepts visually, we will construct a diagram. The term "disorder" will refer to the disease or condition that we wish to diagnose. The term "test" will refer to the radiologic or other clinical test we wish to perform on the patient to determine whether the disorder is present. "Test results" can be either positive or negative. The definition of positive and negative is determined ahead of time by, for example, setting a range of normal values, or using generally accepted criteria when interpreting images.

The diagram can be thought of as a kind of little machine with one moving part. It consists of a rectangle moving around on a four quadrant axes system.

To begin with, each patient will be represented by a short line segment. If the patient truly has the disorder under consideration, we will use a vertical line segment, and if they do not, we will use a horizontal line segment. These line segments are combined to form a box (Figure 1). All of the patients with the disorder are on the vertical side of the box, and all of the patients without the disorder are on the horizontal side of the box.
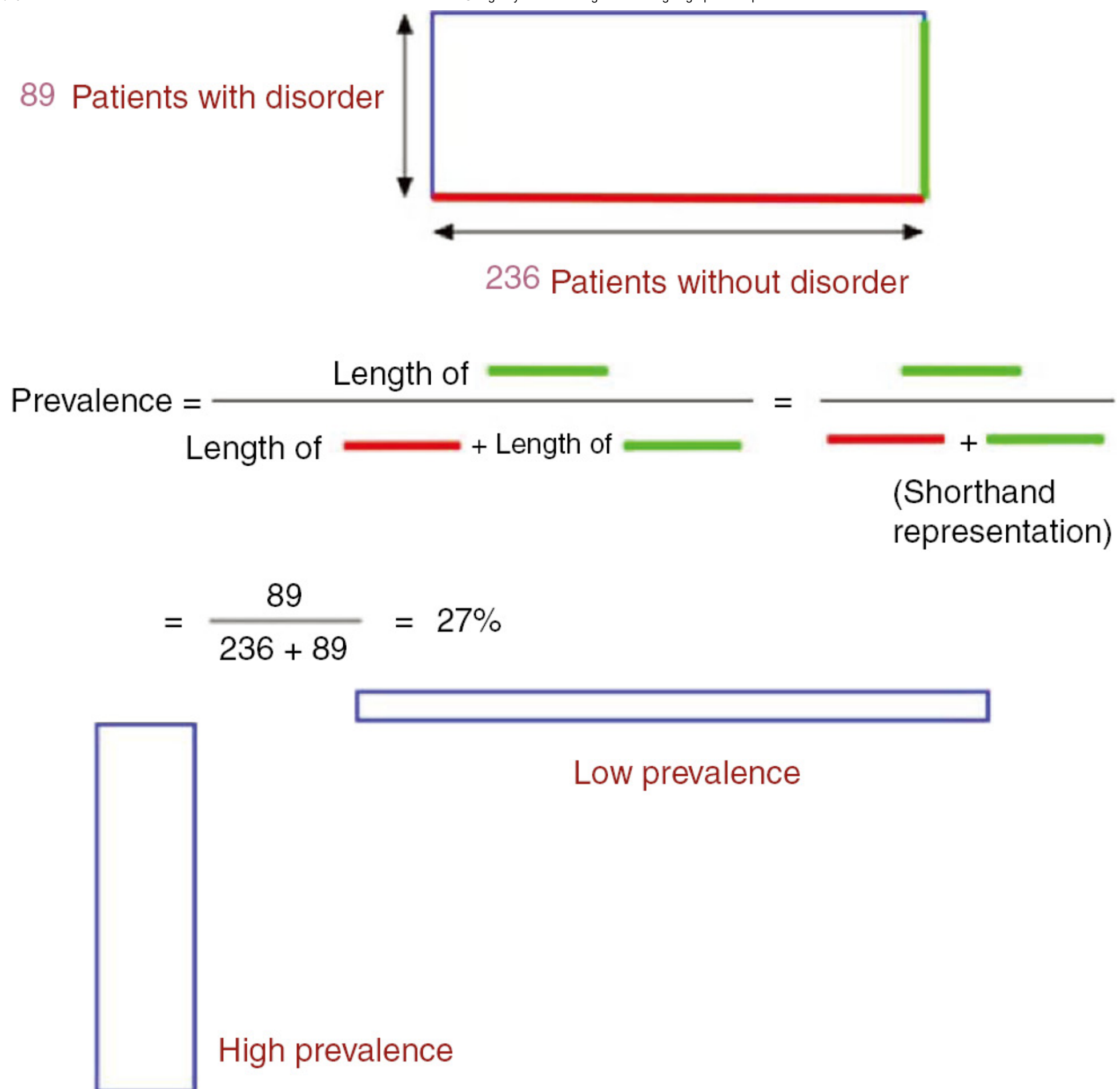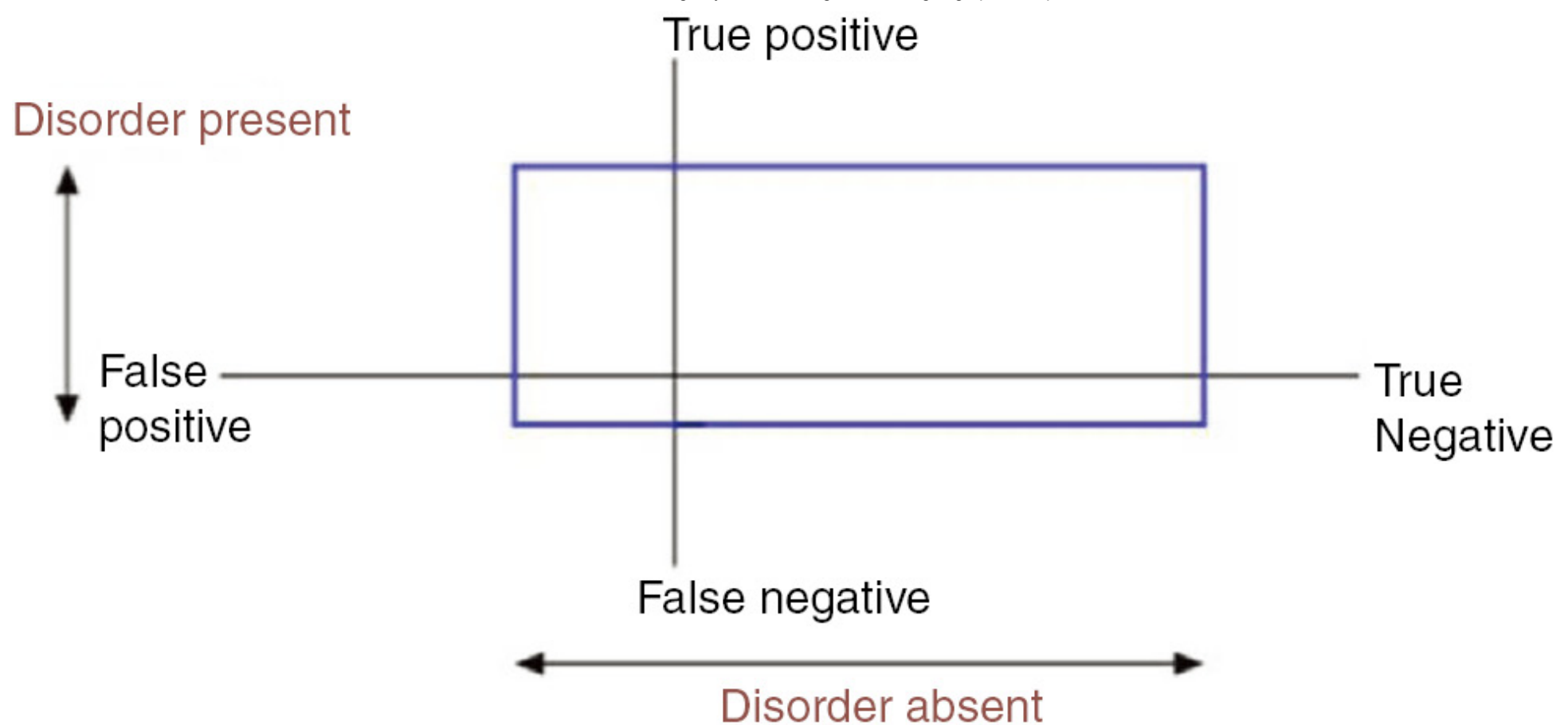
Figure 1:

The patient box is constructed by making the vertical side proportional to the number of patients with the disorder under consideration.

The horizontal side is proportional to the number of patients without the disorder. In published research and in clinical application, this prevalence, or pretest probabilities, can vary widely.

If relatively few patients have the disorder, the box will be low and wide; if many patients have the disorder, the box will be high and narrow. Therefore, the shape of the box tells us at a glance something important about our population. When a study population is considered, the pretest probability is equivalent to the prevalence, that is, the proportion of the population that has the disorder. The prevalence is expressed as the number of patients with the disorder divided by the total number of patients.

Next, we construct a set of axes (Figure 2). The distance along each direction is in units of "number of patients". We place the box on the axes. For now, just think of it as being dropped somewhere in the diagram. The one constraint is that the box must always contain the origin point (0, 0). Now think about sliding the box around on the graph; the box itself does not change shape or size, it simply can slide around to assume different positions on the axes.

Figure 2:

The patient box (blue) is placed onto an axes system with each hemiaxis as marked.

Patients with the disorder are represented on the vertical axes and normal patients on the horizontal axes. The right upper quadrant contains all of the correct test results. The other three quadrants represent error of various kinds.

So what do the axes mean? Because we know that the vertical side of the box represents the patients *with* the disorder, the vertical axis of the graph must refer somehow to those patients. We declare that the distance above zero represents the number of patients with the disorder who also have a positive test result (true positives). It follows that the distance below zero represents the number of patients with the disorder but with a negative test result (false negatives). Thus, based on the results of the test under consideration, we know what the vertical position of the box should be.

The horizontal position of the box is determined in an analogous manner. As defined above, the horizontal side of the box represents the patients *without* the disorder. We declare that the distance to the left of zero represents the number of patients without the disorder who also have a positive test result (false positives). It follows that the distance to the right of zero represents the number of patients without the disorder who also have a negative test result (true negatives).

This is the basic form of the diagram [4, 5]. Every part of the following discussion will refer back to this form, so a thorough grasp of it is essential. I have called it the two- by-two diagram because it is a graphical representation of the two-by-two table used throughout medical statistics. It is a visual embodiment of a very old concept. An online applet is available for readers to use to make their own diagrams [6].
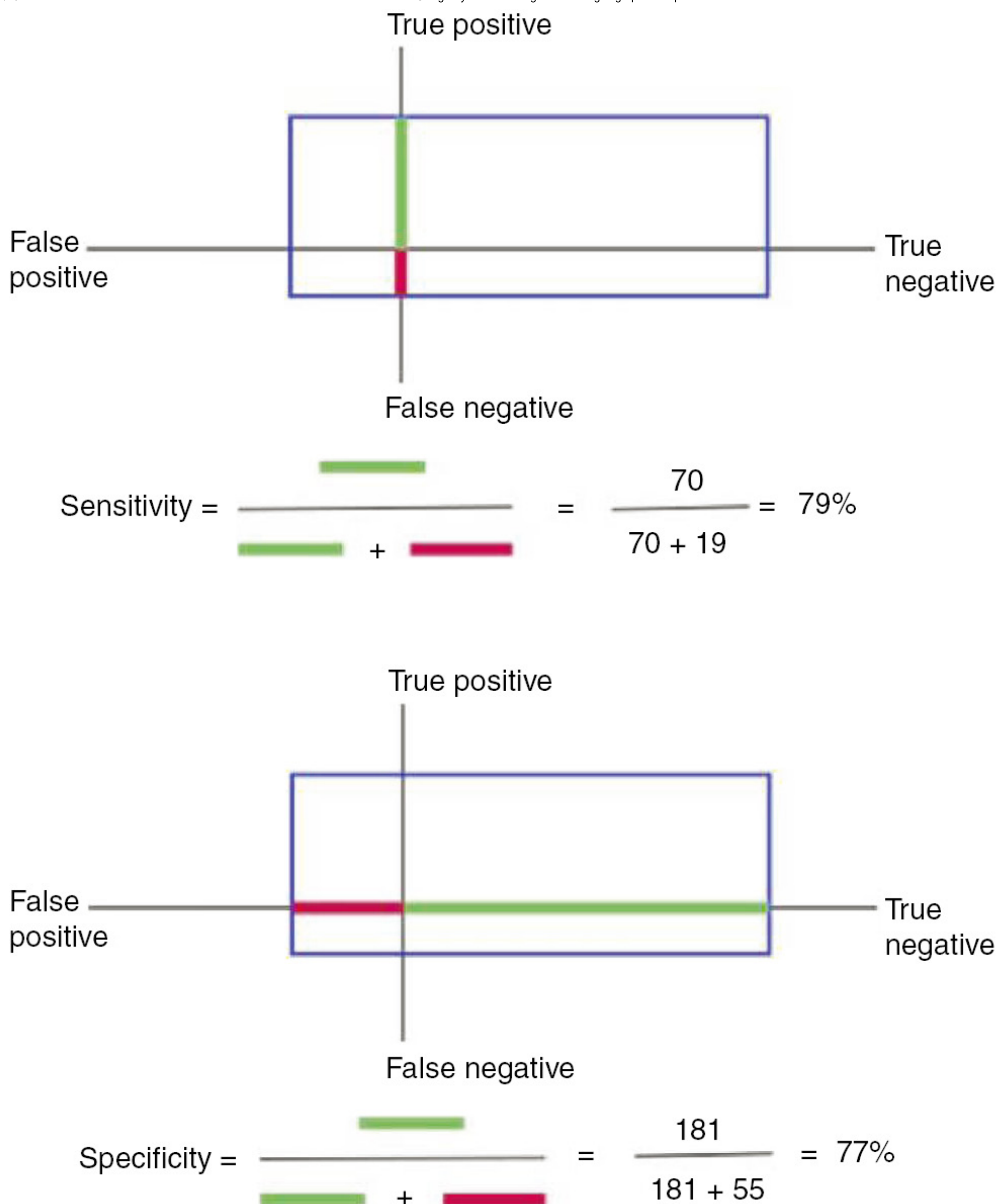
Note that in order to construct the diagram, we have had to make some presuppositions. The truth as to whether the disorder is in fact present must have already been established by some outside reference test that we assume exists but do not further discuss, so whatever limitations arise from that source will be "baked into" our analysis. If the reference test was less than perfect, whatever inaccuracies it contains are carried forward into our own testing, and there is nothing we can do about that after the fact. Therefore, we must take care to select only well-done research studies on which to base our "truth". Also, the definition of "positive" versus "negative" for a test result is needed in those cases where results are not dichotomous, e.g. Doppler velocities and laboratory tests.

It is remarkable how many medical literature articles do not provide enough information to construct the whole diagram [4]. Sometimes this is unavoidable, as when only positive test results lead to the next diagnostic step, which may be invasive or expensive, e.g. biopsy or surgery. In some cases, the information is simply not obtained or not provided.

Sensitivity and specificity graphically defined

Now we need to define sensitivity and specificity. Sensitivity is defined as the proportion (or percentage) of patients *with* the disorder who have a positive test. In the diagram, it is the proportion of the entire vertical dimension of the box that lies above zero (Figure 3 top). It is defined using only the patients with the disorder and has no meaning in regard to patients without the disorder, so the horizontal dimension has no bearing on it. Boxes of many shapes can have the same sensitivity.

$$\text{Sensitivity} = \frac{\rule{3cm}{0pt}}{\rule{2cm}{0pt} + \rule{2cm}{0pt}} = \frac{70}{70 + 19} = 79\%$$

$$\text{Specificity} = \frac{\rule{3cm}{0pt}}{\rule{2cm}{0pt} + \rule{2cm}{0pt}} = \frac{181}{181 + 55} = 77\%$$

Figure 3:

Sensitivity and specificity.

Specificity is defined as the proportion (or percentage) of patients *without* the disorder who have a positive test. In the diagram, it is the proportion of the entire horizontal dimension of the box that lies to the right of zero (Figure 3 bottom). It is defined using only the patients without the disorder and has no meaning in regard to patients with the disorder, so the vertical dimension has no bearing on it. Boxes of many shapes can have the same specificity.

Therefore, as the box slides upward, the sensitivity improves. As it slides rightward, the specificity improves. If the box can reach the right upper quadrant exclusively, it is a perfect test: 100% sensitivity and 100% specificity. Almost no real tests can actually achieve this.

## Side note: sensitivity and specificity are linked

Even at this stage of our explanation, you can see that sliding the box around means that various values must change in an interlinked way. You cannot decrease the number of false positives without increasing the number of true negatives for example. This makes sense.

However, for real tests, it turns out that the movement of the box is further constrained. Almost always, sliding the box to the right is accompanied by a strong tendency for the box to simultaneously slide downwards. That means that although you will reduce the number of false positives, you will also increase the number of false negatives. We account for this behavior by adding a curved line into the left upper quadrant and require that the upper left corner of the box be attached to this line, and stay on the line at all times, thereby constraining the possible positions of the box. This will be called the "test trajectory" (Figure 4A). Its precise shape depends on characteristics of the test and is closely related to the more familiar receiver operating curves.

Setting test cutoff value
to increase specificity
decreases sensitivity
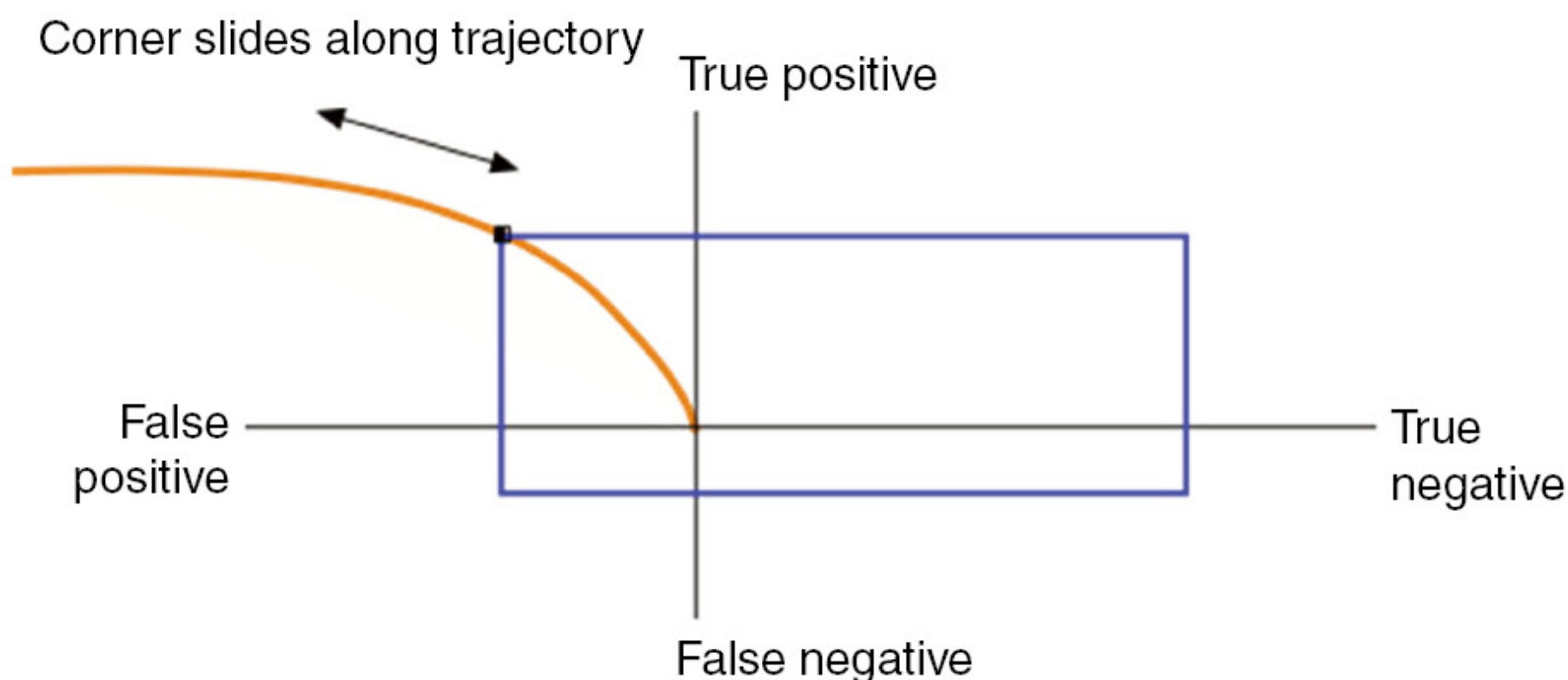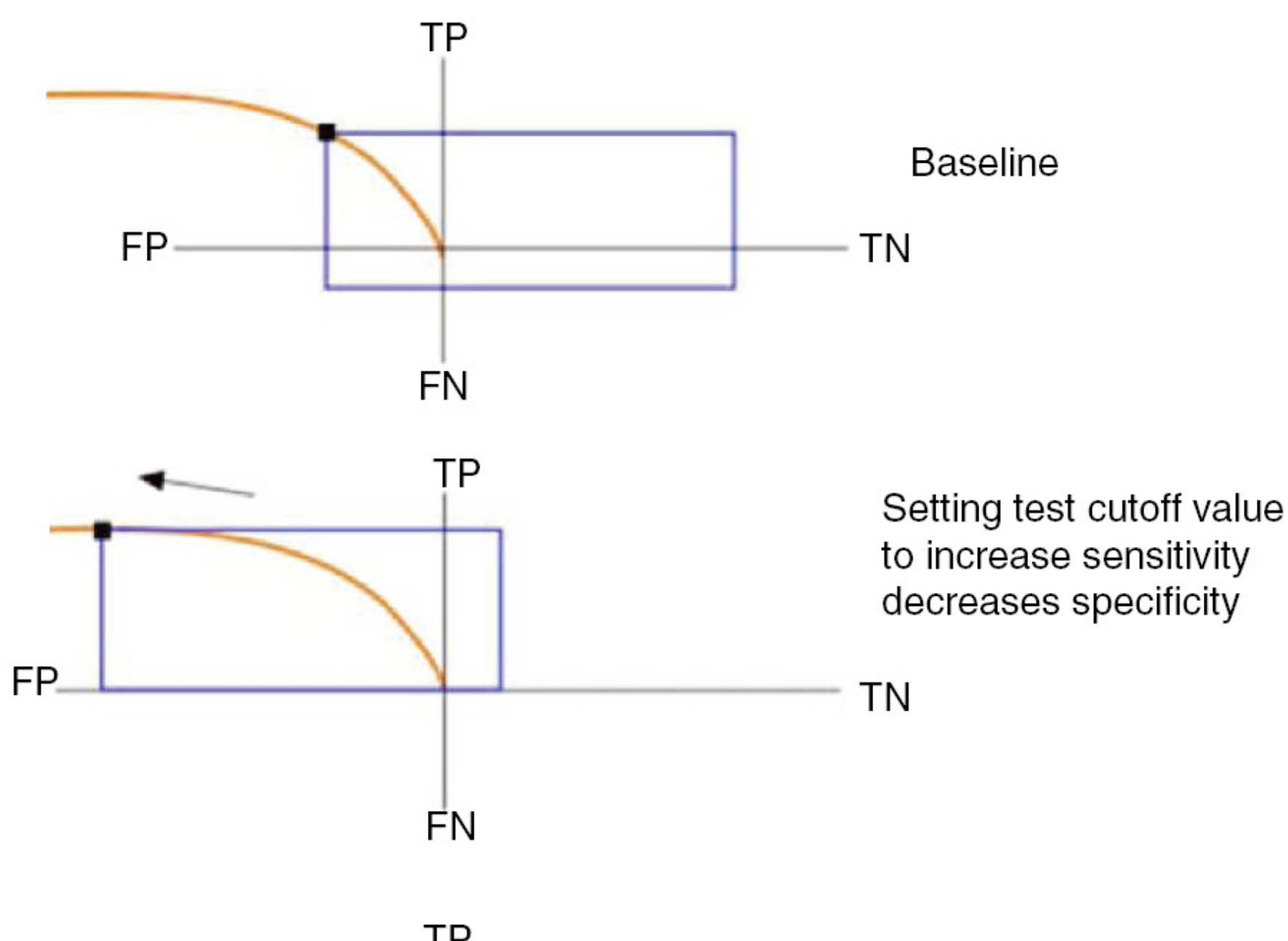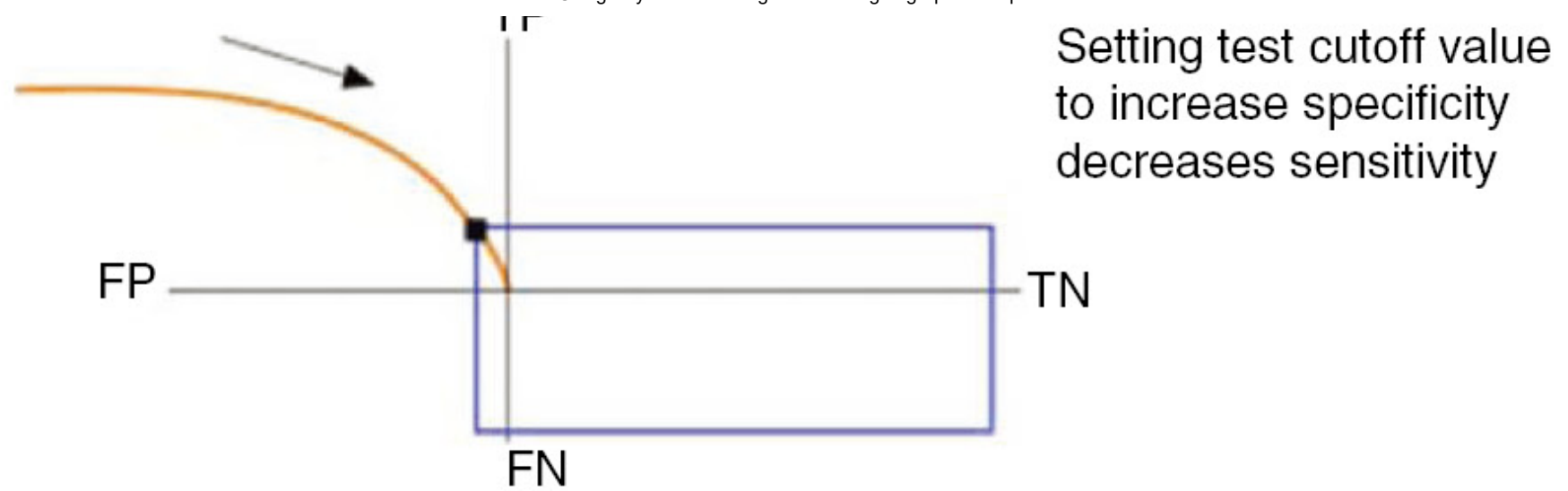
Figure 4:

(A) The patient box (blue) is confined in its movement by the test trajectory. The trajectory is analogous (and mathematically related to) the receiver operating characteristic curve. (B) As the patient box (blue) slides along the test trajectory, sensitivity and specificity change in tandem.

Remember that the position along the test trajectory is determined by adjusting the cutoff criteria by which a test is defined as either "positive" or "negative".

Sensitivity and specificity thus almost always move in opposite directions as the test threshold is changed (Figure 4B).

Our goal: posttest probabilities (positive and negative predictive values)

Now we can find the positive and negative predictive values, i.e. the posttest probabilities. Positive predictive value is defined as the proportion (or percentage) of patients with a *positive* test result who do not have the disorder. In the diagram, it is the proportion of all positive results that lie above zero (Figure 5 top). It is defined using only the patients with a positive test result and has no meaning in regard to patients with a negative result.
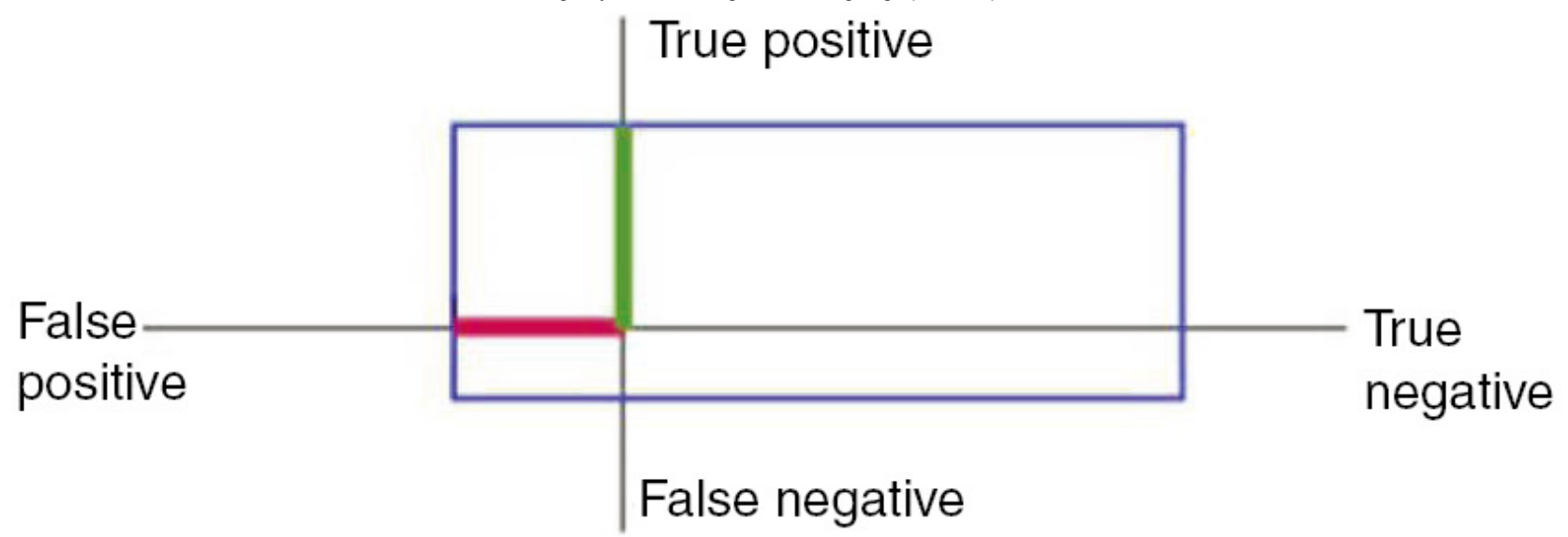
$$\text{Positive predictive value} = \frac{\rule{3cm}{0.3cm}}{\rule{3cm}{0.3cm} + \rule{3cm}{0.3cm}} = \frac{70}{70 + 55} = 56\%$$
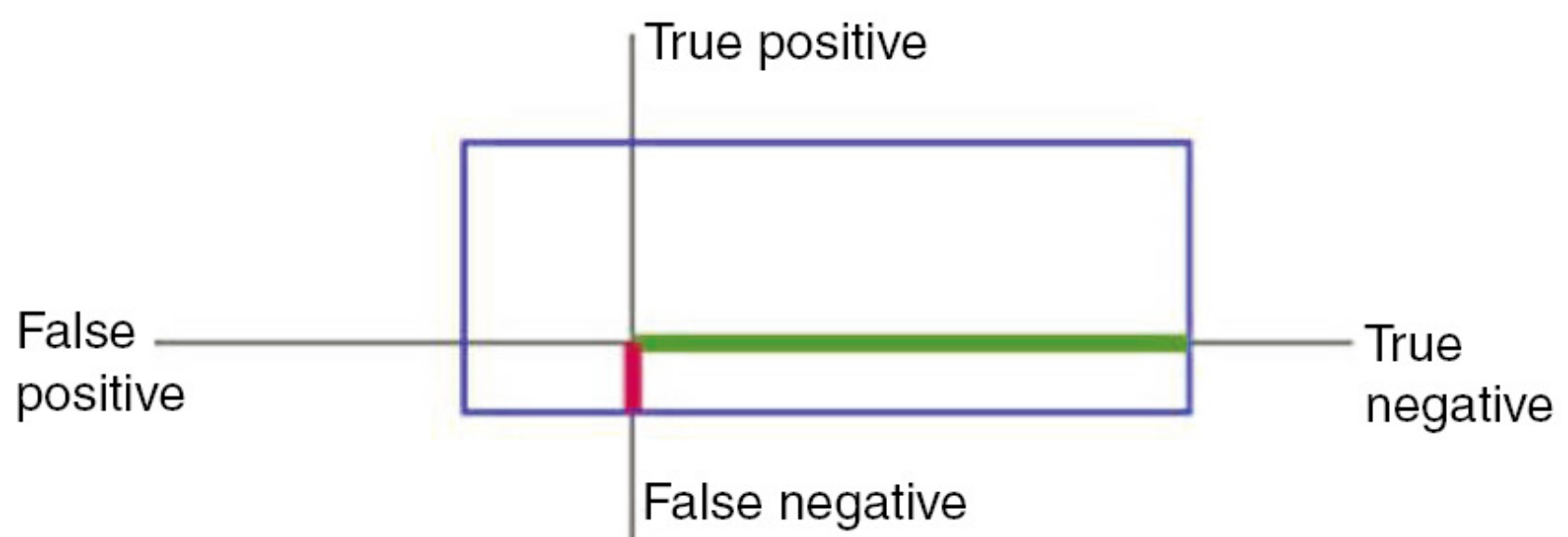
$$\text{Negative predictive value} = \frac{\rule{3cm}{0.3cm}}{\rule{3cm}{0.3cm} + \rule{3cm}{0.3cm}} = \frac{181}{181 + 19} = 90.5\%$$

Figure 5:

Predictive values.

Negative predictive value is defined as the proportion (or percentage) of patients with a *negative* test result who have the disorder. In the diagram, it is the proportion of all negative results that lie to the right of zero (Figure 5 bottom). It is defined using only the patients with a negative test result and has no meaning in regard to patients with a positive result.

Remember that when we talk about sensitivity and specificity, we are given the underlying truth and try to predict the subsequent test results. In real life, we do not already know the truth. By contrast, when we talk about predictive values, we are given the test results and try to predict the underlying truth. It is easy to confuse these two things. However, you can see from the diagram (Figures 3 and 5) how they are different.

Getting to the goal: likelihood ratios and odds

To apply Bayes' theorem, we need more than just sensitivity, specificity and test results. We need an additional piece of information: the pretest probability that the target disorder is present. To estimate that, we need to synthesize our clinical judgements with data from the medical literature. We must be sure to rely on studies that reflect a study population similar to that from which our present patient derives. This is a major challenge but cannot be addressed in further detail here [2, 3].

Our (pretest) suspicion that a disorder is present is usually expressed as a probability. How do we use test sensitivity and specificity numbers to convert this into a posttest probability?

In the usual form of Bayes' rule, the probabilities are used directly, and the answer obtained is also in terms of probability. However, when presented this way, the calculation can seem obscure (see Appendix), when in fact the underlying principle is very simple. For the graphical explanation, and indeed for practical applications in the clinic, a different form of the rule is preferable. That form is expressed in terms of odds, not probabilities, and in terms of likelihood ratios, not sensitivities and specificities.

In this section, we will define odds and likelihood ratios, followed by Bayes' rule in the subsequent section.

Odds

Consider the following transformation of probabilities into odds:
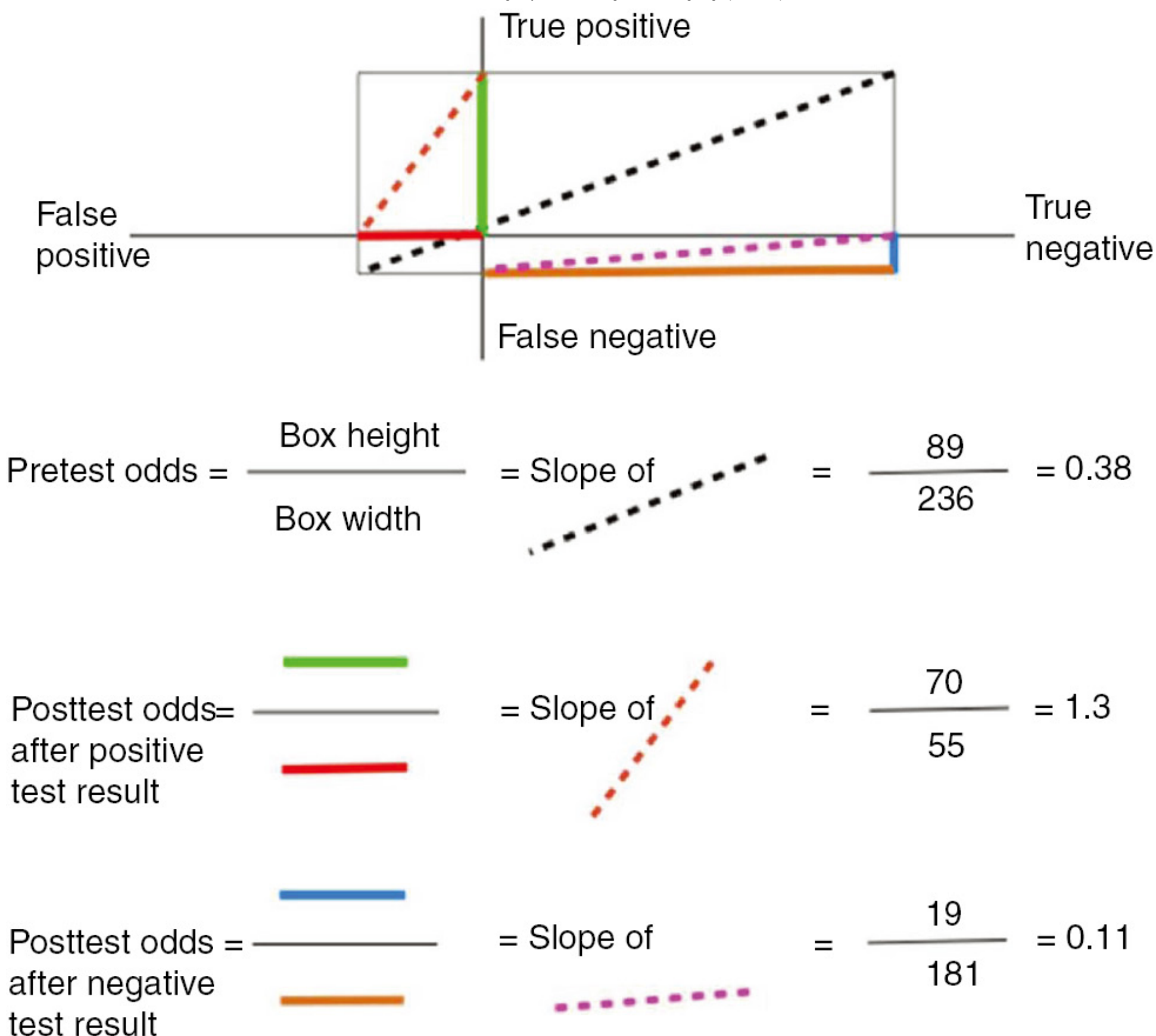
$$\mathrm{Odds\_in\_favor} = \mathrm{probability}/(1 - \mathrm{probability})$$

$$\mathrm{Probability} = \mathrm{odds\_in\_favor}/(1 + \mathrm{odds\_in\_favor})$$

Say we think there is a 25% chance that Mrs. Jones has a pulmonary embolus. Her odds in favor of having an embolus are then 25%/(100−25)%=1/3=0.33. Note that there is an identical amount of information in either probabilities or odds, and they can be interconverted [7].

Odds are not to be confused with odds ratios. Odds are a ratio of frequencies. Odds ratios are a ratio of two odds [4, 7].

Odds appear as slopes in the diagram (Figure 6). Pretest odds equal the number of patients with the disorder divided by the number without the disorder. If the test result is positive, posttest odds in favor of having the disorder equal the number of patients with a true positive test result divided by the number with a false-positive test result. If the test result is negative, posttest odds, again *in favor* of having the diagnosis, equal the number of patients with a false-negative result divided by the number with a true negative result.

$$\text{Pretest odds} = \frac{\text{Box height}}{\text{Box width}} = \text{Slope of} \quad = \frac{89}{236} = 0.38$$

$$\text{Posttest odds} = \frac{\text{(green)}}{\text{(red)}} = \text{Slope of} \quad = \frac{70}{55} = 1.3$$

after positive test result

$$\text{Posttest odds} = \frac{\text{(blue)}}{\text{(orange)}} = \text{Slope of} \quad = \frac{19}{181} = 0.11$$

after negative test result

Figure 6:

Odds are defined as ratios of abnormal to normal patients. Pretest odds are the number of abnormal patients divided by the number of normal patients (first equation). This ratio can be depicted as the slope of the black line. Likewise, if the test is positive, the posttest odds in favor of having the disorder is represented by the slope of the brown line in the left upper quadrant. If the test is negative, the posttest odds are given by the slope of the purple line in the right lower quadrant. As test performance improves, the box moves into the right upper quadrant, and the brown slope increases while the purple slope decreases.

Likelihood ratios

Now consider the following transformation of sensitivity and specificity into positive and negative likelihood ratios:
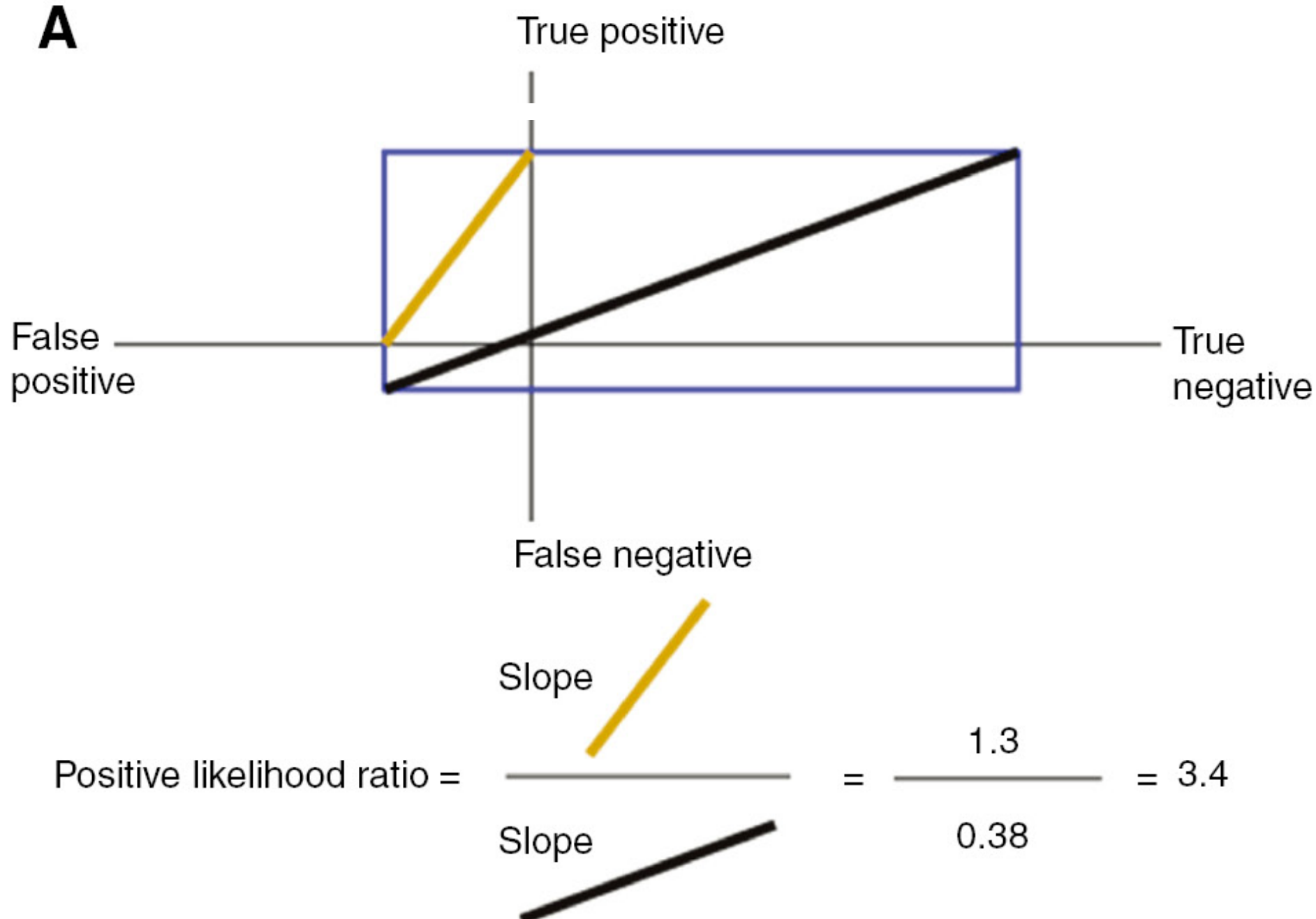
Positive likelihood ratio $\text{LR}+ = \text{sensitivity}/(1 - \text{specificity})$

Negative likelihood ratio $\text{LR}- = (1 - \text{sensitivity})/\text{specificity}$

Just like sensitivity and specificity, likelihood ratios come in pairs. Here too there is an identical amount of information in either set of concepts. If we know the sensitivity and specificity, we can easily calculate the likelihood ratios, and vice versa. The positive likelihood ratio (LR+) refers only to positive test results. The negative likelihood ratio (LR−) refers only to negative test results.

This recasting into positive and negative likelihood ratios has the advantage of moving the focus away from what is known from the reference standard (the underlying truth, unknown to us in a clinical scenario) towards what might be knowable from the test.

In the diagram, likelihood ratios are represented by the ratio of two slopes. Each slope is an odds value, as explained above. Positive likelihood ratio is the slope of the diagonal of the left upper quadrant of the box (post– test odds in favor of having the disorder given a positive test) divided by the slope of the diagonal of the entire box (pretest odds, Figure 7A). Negative likelihood ratio is the slope of the right lower quadrant of the box (post– test odds in favor of having the disorder given a negative test) divided by the slope of the diagonal of the entire box (Figure 7B).
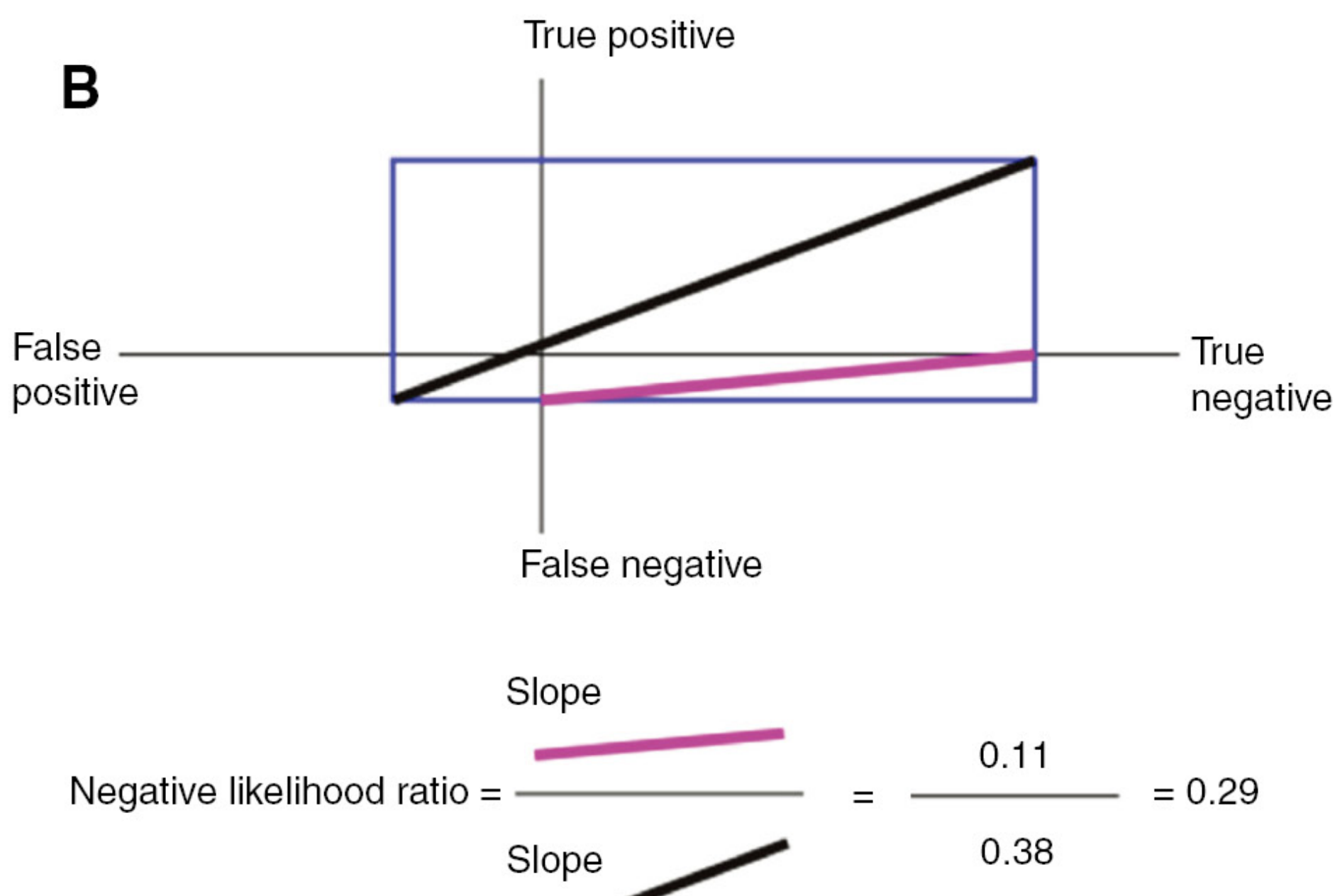
Figure 7:

(A) The positive likelihood ratio is the ratio of two odds, that is, the odds in favor of having the disorder once a positive test result is known (slope of brown), divided by the odds in favor of having the disorder before the test was performed (slope of black). Conventionally, a value of 10 or greater is considered "good". (B) The negative likelihood ratio is the ratio of two odds, that is, the odds in favor of having the disorder once a negative test result is known (slope of purple), divided by the odds in favor of having the disorder before the test was performed (slope of black). Conventionally, a value of 0.1 or less is considered "good".

Restated, the positive likelihood ratio is the posttest odds of a positive test result divided by the pretest odds that the disorder is present. The negative likelihood ratio is the posttest odds of a negative test result divided by the pretest odds that the disorder is present.

When the LR+ is high (good), the slope based on only the positive test results (left upper quadrant) is much steeper than the slope of the entire box. When the LR− is low (also good), the slope based on only the negative test results (right lower quadrant) is much less steep than the slope of the entire box.

Think about moving the box around and how that would influence the slopes and therefore the likelihood ratios. It is easy to visualize how these slopes must change as we move the box upward or downward, right or left. Simultaneously, we see how the sensitivity and specificity are changing.

As we move the box into the right upper quadrant (the "best" quadrant), the slope for the left upper quadrant diagonal gets steeper until eventually it becomes infinite.

At the same time, the slope for the right lower quadrant diagonal becomes smaller until eventually it is zero.

Bayes' rule

According to Bayes' rule (or theorem), the posttest odds equal the pretest odds multiplied by the likelihood ratio. This is just stating what has already been explained above.

$$\text{likelihood ratio} = \text{pretest odds}/\text{posttest odds}$$

or rearranging

$$\text{posttest odds} = \text{pretest odds} \times \text{likelihood ratio}$$

If the test result is positive, use the positive likelihood ratio for this calculation. If the test result is negative, use the negative likelihood ratio instead. In both scenarios, the odds refer to the odds *in favor* of the target disorder being *present*.

For our purposes, there is no more to Bayes' rule than this. The equation as expressed in probabilities is given in the appendix.

In practice, tests with positive likelihood ratios >10 and negative likelihood ratios <0.1 are considered "good".

So LR+=10 means that, if the test result is positive, the posttest odds of having the disorder are 10 times higher than the pretest odds. That seems pretty good, but keep in mind that if the pretest odds are very low then the posttest odds might still be only fair at best. For example, the pretest odds of having cancer on a screening mammogram are about 6/1000. The positive likelihood ratio is about 14, so the posttest odds are 84/1000. Whether the positive patients have cancer remains unresolved, and they must go on to further testing and/or biopsy. Here again we see the importance of prevalence when interpreting test results.

LR−=0.1 means that the posttest odds of having the disorder are one tenth as high as the pretest odds. So if the pretest odds of a deep venous thrombosis are 1/5, and the D dimer is negative (LR−=0.05), the posttest odds are only 1/100, helping to rule out the diagnosis.

Imagine moving the box towards the right. As it moves to the right, it tends to also move downward. The trajectory constrains the possible path. The slope for LR+ increases towards infinity (good). However, the slope for the LR− also increases (not good). Trade-offs are inevitable.

Likelihood ratios are a function only of sensitivity and specificity. They do not contain any more information than sensitivity and specificity alone but simply recombine that information into a different form. Therefore, likelihood ratios are subject to the same reservations and limitations. For example, because sensitivity and specificity depend on patient spectrum and on cut points, so do likelihood ratios.

Conclusions

In medicine, Bayes' theorem is a mathematical statement of how a test with known sensitivity and specificity changes the probability that a disorder is present. The two-by-two diagram is a visual way to think about this concept. If the theorem were different, the diagram would have to be different, and if the diagram were much different, it would not embody the theorem. Given the same quantitative inputs to the diagram and the theorem, the outputs are the same. An online applet is available for readers to use to make their own diagrams (6).

Likelihood ratios can be used in clinical practice given some effort. Historically, we have discussed tests in terms of sensitivity and specificity rather than positive and negative likelihood ratios, and in terms of probabilities rather than odds. There is no reason why the second pair of concepts could not replace the first, but that is not the world we live in. Unfortunately, this makes application of Bayes' theorem a bit unwieldy because of the need to convert probabilities to odds and back again. A simple calculator suffices, but smart phone apps exist that simplify matters [8, 9].

At the moment, there is strong interest in "clinical decision support" for clinicians ordering imaging tests. This consists of a software interface that is supposed to help the clinician decide whether a given clinical problem can be effectively addressed by a given test. One of the most popular products is based on the "appropriateness criteria" from the American College of Radiology [10]. However, the ACR criteria do not explicitly list likelihood ratios. Instead, a presenting symptom and/or sign is listed (implying the underlying pretest probability of the disorder) followed by an "appropriate use" score from 1 to 9 for each of several candidate tests (implying likelihood ratios).

However, an optimal estimate of pretest probability must be based on the full, sometimes subtle, clinical picture. The tool does not explicitly list the probability presumed by the authors for each given indication, so the clinician has no basis on which to adjust that part of the analysis. Also, there is no distinction made between positive and negative likelihoods. Does a test get a low rating because it is insensitive or nonspecific? The different implications of a negative versus a positive test result are not apparent. In short, the present format does not lend itself to the kind of analysis described in the present paper. The criteria also do not consider the order in which tests are performed, which can markedly affect the overall accuracy of the diagnostic endeavor.

Although sensitivity and specificity appear everywhere in the radiology literature, likelihood ratios are much less commonly cited (though easily calculated). Clinicians need reliable lists of likelihood ratios for radiologic tests under various conditions, updated periodically [11]. In particular, it would be useful to flag exams that have high utility, i.e. those with LR+>10 or LR−<=0.1, with additional information on patient spectrum and cut points. A list of particularly poor tests would also be helpful as well as a list of those often misused.

The existing appropriate use criteria could be revised by the professional consensus panels to be expressed in these terms. The result might be an exceptionally useful clinical tool, rather than a fixed checklist/ranking system.

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: None declared.

Employment or leadership: None declared.

Honorarium: None declared.

Competing interests: The funding organization(s) played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

Appendix

Bayes' theorem is simply the mathematical statement that connects the pretest probability to the posttest probability. The theorem is another way of stating the same relationships illustrated in the diagrams using likelihood ratios above. The following explanation is provided for completeness sake for interested readers [12]. Reference [6] provides a link to a simple applet to draw the two by two diagram.

Mathematically, the theorem as presented using probabilities is

$$P(D|A) = (P(A|D) \times P(D))/P(A)$$

where test result A can be either positive or negative and

P(D|A) is the probability of having the disorder D given test result A

P(A|D) is the probability of having the test result A given that the disorder D is present

P(D) is the pretest probability of having the disorder D

P(A) is the overall probability of the test result A.

References

1.	Stein PD, Hull RD, Patel KC, Olson RE, Ghali WA, Brant R, et al. D-dimer for the exclusion of acute venous thrombosis and pulmonary embolism – a systematic review. Ann Intern Med 2004;140:589–602. [Search in Google Scholar](#)

2.	Willis BH. Spectrum bias – why clinicians need to be cautious when applying diagnostic test studies. Fam Pract 2008;25:390–6. [Search in Google Scholar](#)

3.	Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. N Engl J Med 1978;299:926–30. [Search in Google Scholar](#)

4.	Johnson KM, Johnson BK. A visual presentation of statistical concepts in diagnostic testing: the two by two diagram. Am J Roentgenol 2014;203:W14–20. [Search in Google Scholar](#)

5.	Johnson KM. The two by two diagram: a graphical truth table. J Clin Epidemiol 1999;52:1073–82. [Search in Google Scholar](#)

6.	Two by Two Diagram display and calculator. http://54.221.244.58: 3838/Truth/. [Search in Google Scholar](#)

7.	Probability versus odds. http://stats.seandolinar.com/statistics-probability-vs-odds/. Accessed 16 Feb 2017. [Search in Google Scholar](#)

8.	Evidence Based Medicine calculator – Knowledge Translation Clearinghouse, Canadian Institute of Health Research. http://ktclearinghouse.ca/cebm/practise/ca/calculators/statscalc. Accessed 16 Feb 2017. [Search in Google Scholar](#)

9.	Example of smart phone app for likelihood ratios. https://itunes.apple.com/us/app/dxlogic/id943680508?mt=8. [Search in Google Scholar](#)

10.	ACR appropriateness criteria. https://www.acr.org/Quality-Safety/Appropriateness-Criteria. Accessed 11 Mar 2017. [Search in Google Scholar](#)

11.	Details on Likelihood Ratios – Knowledge Translation Clearinghouse, Canadian Institute of Health Research. http://ktclearinghouse.ca/cebm/glossary/lr. Accessed 23 Feb 2017. [Search in Google Scholar](#)

12.	Moye LA. Elementary Bayesian Biostatistics. New York: Chapman and Hall/CRC Biostatistics Series, 2007. [Search in Google Scholar](#)

Download article (PDF)

From the journal

Diagnosis
Volume 4 Issue 3

Journal and Issue

Search journal

○ This issue   ○ All issues

Articles in the same Issue

[Frontmatter](#)

[Improving diagnosis in radiology – progress and proposals](#)

[Improving diagnosis in health care: perspectives from the American College of Radiology](#)

[The role of radiology in diagnostic error: a medical malpractice claims review](#)

[Medical errors, malpractice, and defensive medicine: an ill-fated triad](#)

[Perceptual errors in pediatric radiology](#)

[256 Shades of gray: uncertainty and diagnostic error in radiology](#)

[Using Bayes' rule in diagnostic testing: a graphical explanation](#)

[X-ray art images: objects from nature](#)

[Assigning responsibility to close the loop on radiology test results](#)

[See more...](#)

Contact us

Customer Service

Human Resources

Press

Contacts for authors

Career

How to join us

Current Vacancies

Working at De Gruyter

Open Access

Articles

Books

Funding & Support

For Authors

Publish your book

Publish your journal article

Abstracting & Indexing

For Libraries & Trade Partners

Electronic Journals

Ebooks

Databases & Online Reference

Metadata

Our Partner Publishers

Rights & Permissons

Repository Policy

Free Access Policy

About De Gruyter

De Gruyter Foundation

Our locations

Help/FAQ

Privacy Policy

Terms & Conditions

Legal Notice