



5 Clinical Decision-Making in Cardiology

JOHN E. BRUSH JR. AND HARLAN M. KRUMHOLZ

DIAGNOSTIC DECISIONS, 53

THERAPEUTIC DECISIONS, 57

DECIDING WHEN TO CHANGE CLINICAL PRACTICE BASED ON NEW CLINICAL RESEARCH FINDINGS, 59

SHARED DECISION-MAKING, 59

MONITORING THE QUALITY OF CLINICAL DECISIONS, 60

System 1 and System 2 Thinking, 60

TEACHING CLINICAL REASONING, 60

CONCLUSION, 61

CLASSIC REFERENCES, 61

REFERENCES, 61

Medicine is an information science. Information is being produced at an unprecedented rate and is readily accessible using electronic searches and hand-held devices, making skills to parse and use the appropriate information ever more important. Memorization of medical facts is less a necessity, while processing knowledge and critical thinking remain essential for high-value medical care. Clinical decisions and recommendations are critical features of medicine and, in the midst of a rapid expansion of medical knowledge, have never been more challenging. This chapter summarizes core competencies for clinical reasoning that should be mastered by expert practicing cardiologists.

Excellent clinical decisions require a command of medical knowledge and a deep understanding of individual patients, including their preferences and goals. Good decisions take into account the limits of knowledge, uncertainty in measurements, and the play of chance.¹⁻³ Clinical reasoning is informed by both experiential and formal knowledge learned through years of practice and study.⁴ The translation of medical knowledge into good patient-centered decisions is a key goal of clinical reasoning and is the hallmark of an expert clinician.

Clinical reasoning is often guided by simplified rules. Early in training, physicians are taught how to recognize specific clusters of signs and symptoms, place patients in diagnostic categories, and follow the rules that apply to those categories. For example, patients with particular findings might be labeled as having acute myocardial infarction (AMI), which would trigger treatment based on studies showing benefit from aspirin and beta-blocking agents. In this context, algorithmic tools are often used to direct actions. For example, guidelines recommend that a patient with a low ejection fraction should be considered for an automated implantable defibrillator, but only after considering the etiology of the systolic dysfunction and the timeframe of the disorder.

These rule-based algorithms are not intended to force actions, but to guide decisions. The best clinicians know when adherence to such algorithms is proper and when exceptions, based on the patient's particular situation or preferences, can lead to divergence from these algorithms. Divergence from guidelines may be appropriate, but requires adequate justification, documentation, and transparency.

Most of medical decision-making, however, lies outside of simple algorithms and requires judgment. There are two major settings, related to diagnosis and treatment, where clinical reasoning is critical.

First, there are decisions about classifying an individual who presents with symptoms or signs of disease into the proper diagnostic category. Book chapters and other reference materials are usually organized according to categories, such as a medical diagnosis. The chapter informs the reader about how a particular condition, such as aortic stenosis, might manifest. These labels are useful for clustering patients by common disease mechanisms, prognosis, and responses to therapeutic strategies.

But patients often do not present to medical attention fitting perfectly into pre-specified general diagnostic categories. They seek attention for symptoms, which requires the clinician to reverse the order of a typical textbook and to work inductively from a patient's signs and symptoms toward a diagnostic label before a therapeutic plan can

be developed. For a patient with dyspnea on exertion and a systolic murmur, aortic stenosis is a possibility, but the diagnosis is not conclusive without further testing. In some cases, uncertainty persists. About a third of patients labeled with a principal discharge diagnosis of heart failure also receive treatments for other causes of dyspnea such as pneumonia or chronic obstructive pulmonary disease.⁵ This is the reality of current practice.

Second, there are decisions about treatments. These decisions are also challenging because they involve weighing risks and benefits, speculating about estimates for these parameters, and aligning choices with the preferences of the patient. The likelihood of benefit is often probabilistic, as people are pursuing strategies to reduce risk without knowing whether they themselves will benefit. These decisions can occur in prevention, which addresses whether to intervene in the interest of preventing future health problems, based on an estimate of prognosis. In this setting, the risks and costs occur immediately while the benefit is anticipated to be in the future. These decisions can also involve treatments to address symptoms as well as reduce the immediate risk for someone with acute or chronic disease.

Risk stratification is an important application of probability and is often used to estimate patient risk and assist in decision-making. This approach generally uses the results of statistical models that have identified prognostic factors and incorporated them into a tool that may assist clinicians. In recent years, many tools have been developed to assist in the rapid assessment of patients.

Recent decades have witnessed the emergence of cognitive psychology, a branch of psychology focused on how people make decisions.⁶ The field demonstrated that people frequently develop useful reasoning shortcuts to circumvent the need to explicitly calculate probabilities, but these shortcuts come with biases that can lead decision-makers to deviate from the rules of logic and probability in predictable ways. Thus, a good understanding of clinical reasoning requires knowledge about logic and probability as well as cognitive psychology.

Cognitive psychologists have demonstrated how people often rely on intuition to make decisions in uncertain settings.^{6,7} For cognitive psychologists, intuition is not merely guessing, but has a specialized meaning. The cognitive psychologist Herbert Simon described intuition by stating: "the situation has provided a cue; this cue has given the expert access to information stored in memory, and the information provides the answer. Intuition is nothing more and nothing less than recognition" (see Classic References). Expert clinicians learn to use intuition to recognize diagnoses and make clinical decisions. They learn to calibrate their intuitive judgments using scientific evidence and clinical experience. They may also be susceptible to cognitive biases that are associated with such decision-making.

DIAGNOSTIC DECISIONS

Patients often present with descriptions of symptoms such as chest pain. Cues are scattered, like pieces of a jigsaw puzzle. Clinicians, like all





decision-makers, often use mental shortcuts called heuristics to organize cues and to turn an unstructured problem into a set of structured decisions.^{3,7} They are taught to collect the cues of an unstructured clinical problem by using an organized history and physical examination.⁸ When experts take a history, they use a process known as early hypothesis generation to develop a list of 3 to 5 possible diagnoses very early in the process (see Classic References). This enables the questioning to become more direct and the clinician to become more engaged in the fact-finding exercise.

Studies show that the mechanism of diagnostic hypothesis generation varies, depending on the stage of training.⁸ Novice practitioners who lack clinical experience use causal reasoning, which tends to be slow and less accurate. As trainees gain experience, knowledge about diagnoses becomes encapsulated into illness scripts. An illness script is a schema or map that integrates conceptual information regarding a disease and links the concepts with case experience. As physicians gain further experience, they accrue experiential knowledge. One theory is that diagnostic experience is remembered through disease prototypes, which describe the typical features of a disease. Another theory is that experiential knowledge is remembered as specific instances called exemplars, which are memories of prior experiences that have been categorized and stored in long-term memory. With experience, a clinician accumulates exemplars that are automatically retrievable and represented in memory in a fashion that is unique to that clinician and not generalizable among clinicians. Memories of exemplars give the expert an intuitive sense of both the base rates for particular diagnostic categories and the relative frequencies of features for a diagnostic category.

Because clinicians start the diagnostic process by intuitively recognizing familiar phenotypes stored in memory as exemplars, it becomes important to study how symptoms combine in individuals as unique symptom phenotypes. A recent study showed wide variation in symptom phenotypes among patients with AMI, which may have important implications on how we teach learners to recognize a diagnosis.⁹ The study also showed that women exhibited significantly more unique symptom phenotypes than men. Greater phenotypic variation could lead to more missed diagnoses and this is a promising area for further research.

After collecting, sorting, and organizing clinical data, clinicians often use a problem list as a tool to list, group, and prioritize clinical findings. With additional clinical information, a problem statement can be defined more specifically. For example, shortness of breath may be an initial problem statement that is replaced by acute systolic heart failure, as further clinical information leads to a more refined problem statement that moves from symptom to diagnosis. They then use a differential diagnosis to expand the list of possibilities to avoid premature closure of the search for the true diagnosis. This step-by-step process enables the clinician to formulate a set of hypothetical diagnostic possibilities, which can then be tested using iterative hypothesis testing. Iterative hypothesis testing allows the clinician to narrow the list of possible diagnoses and focus on the most plausible hypothesis.¹⁻³

UNDERSTANDING PROBABILITY

Understanding probability is essential for good clinical decision-making.¹⁻³ Probability can be estimated for outcomes that are measured as continuous or categorical variables, as shown in Figure 5.1. The figure shows how probability of an outcome or event is distributed across a range of possibilities. For example, a laboratory test might be measured in a population of patients resulting in a distribution in which most patients are distributed to the middle of the range of possibilities and fewer scatter to the edges of the range, shown in the probability density curve in the left panel of Figure 5.1. The probability of categories or discrete variables can also be measured, as shown in the probability distribution graph in the right panel of Figure 5.1. If all of the diagnostic possibilities are mutually exclusive and collectively exhaustive, the probability of all of the possibilities will add up to 1, as shown by the red cumulative probability curves in Figure 5.1. Understanding cumulative probability is important for understanding sensitivity and specificity, as discussed below.

To test a diagnostic hypothesis, clinicians use conditional probability, which is the probability that something will happen, on the condition that something else happened. Conditional probability can inform the probability of a diagnosis, on the condition of some new information such as a positive test result. Bayesian reasoning is a mental process that allows clinicians to modify their perceptions by considering prior knowledge and updating that knowledge with new and evolving evidence. It enables formation of a probability estimate and revision of that estimate based on new information using conditional probability. For example, one might ask, what is the probability of coronary artery disease in a patient, given a positive stress echocardiogram? What is the probability of pulmonary embolus, given a negative D-dimer test? What is the probability of an acute coronary syndrome, given an abnormal troponin test? The post-test probability depends on a prior estimate of the probability for that particular patient, combined with the strength of the test result. Probability theory helps the clinician understand the question and calculate the answer.

Bayesian reasoning adds mathematical rigor to clinical thinking and requires both a prior estimate of probability and an estimate of the strength of a test result. Prior estimates can come from a clinician's own experience, or published data on the prevalence of a disease. A classic paper by Diamond and Forrester provides estimates of the prevalence of coronary artery disease in patients depending on age, sex and symptom features, for example (see Classic References). This type of observational research can be used to provide the prior probabilities that are needed for Bayesian reasoning.

Understanding probability is essential to interpreting laboratory tests. A laboratory test might be measured in a population of presumably normal individuals to determine a distribution and to define a normal range, shown in the probability density curve in the left panel of Figure 5.2. A normal range is commonly defined as the inner 95% cumulative probability and the abnormal range is defined as values falling outside of the normal range as shown.

Another way of defining a test result is by measuring the test result in a group of subjects who are defined as normal and abnormal by another independent "gold standard" test, as shown in the right panel of Figure 5.2. Typically, subjects with and without disease will have test results that are distributed like bell-shaped curves. A line of demarcation can be drawn to define how a new test would separate patients with positive and negative test results. Because there is overlap in subjects

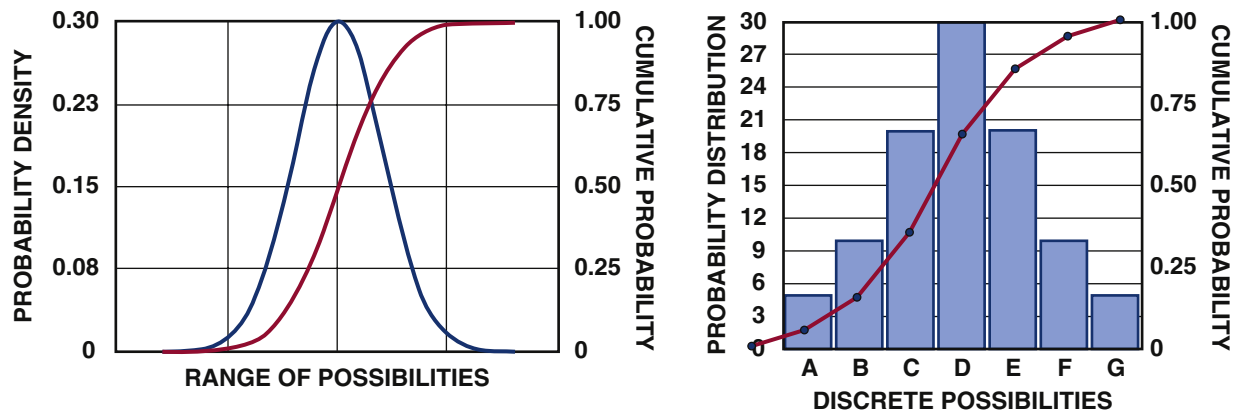


FIGURE 5.1 A probability density curve is shown in the left panel. The blue curve shows the probability of an event (left y-axis), across a range of possibilities (x-axis). A probability distribution is shown in the right panel. The blue columns show the probabilities (left y-axis) of a variety of discrete possibilities (x-axis). In both panels, the cumulative probability across the range of possibilities (x-axis) is shown by the red curves (right y-axis).

with and without disease, there will be false-positive and false-negative test results, as shown.

Understanding how to use clinical testing is essential to good decision-making. The utility of a test result depends, in part, on the operating characteristics of a test, namely, the sensitivity and specificity. They are rates, meaning they are proportions with different units for the numerator and denominator. The terms “true positive rate” (TPR) for sensitivity and “true negative rate” (TNR) for specificity are alternative labels. Patients with and without disease are shown separately in Figure 5.3 to show the cumulative probabilities of a true positive result (sensitivity or the TPR) on the right and of a true negative result (specificity or the TNR) on the left. Sensitivity and specificity are usually shown in a 2 × 2 table but showing the TPR and TNR in Figure 5.3 demonstrates how these rates vary, depending on the location of the line of demarcation between positive and negative test results.

The complementary probability of the TNR is the false-positive rate (FPR), as shown in the top panel of Figure 5.4. Plotting the TPR (sensitivity) of a test on the y-axis and the FPR (1-specificity) on the x-axis creates a plot called a receiver operating curve (ROC), as shown in the bottom panel of Figure 5.4. ROCs are useful for determining the optimal cutoff point for the line of demarcation of a test.

The denominators of sensitivity and specificity are patients with the disease and people without the disease, respectively. In clinical practice, when test results are reported as positive or negative, however, the results are reported using terms with different denominators. A clinician wants to know the probability that a positive test result is truly positive, or the positive predictive value (PPV), and also the probability of disease given a negative test result, which is 1 minus the negative predictive value (NPV). When changing from sensitivity and specificity to the PPV and NPV, the denominators of these rates change, making it difficult for a clinician to estimate these probabilities intuitively. In addition, the PPV and NPV depend not only on the sensitivity and specificity of the test, but also on the prevalence of the target condition in a population of test subjects. The sensitivity and specificity are not fixed, and spectrum bias can result if the test subjects that defined the operating characteristics of the test

are different from the subjects who are subsequently tested.^{2,3} If the operating characteristics of the test are defined in a narrowly defined population (left panel of Fig. 5.5), but the test is used in a broadly defined population and the line of demarcation remains fixed (right panel of Fig. 5.5), the specificity, or TNR, will decrease. This commonly occurs with tests such as troponin testing, where the clinical sensitivity and specificity of the test are defined in a research setting, but the test is used indiscriminately in practice. When used as a general screening test in a broadly defined population, the width of the distribution of the subjects with no disease widens, yet the line of demarcation remains fixed, which decreases the TNR, as shown. This issue has also been shown in genetic testing.¹⁰

In practice, clinicians usually do not formally calculate Bayesian probabilities but, in general, use a heuristic that psychologists call “anchoring and adjusting.”^{3,6} Clinicians estimate a pretest probability (the anchor) and estimate the posttest probability by adjusting the anchor. For a patient with chest pain, for example, the anchor would be an estimate of the pretest probability of coronary artery disease, which would be intuitively adjusted on the basis of new information such as a stress test result to estimate a posttest probability. This is an expedient method for intuitively estimating conditional probability.

There are two potential problems when using this heuristic. One fallacy, called “anchoring,” is when the decision-maker becomes too anchored on the pre-test probability estimate and does not adequately adjust in estimating the post-test probability. The second fallacy is called “base-rate neglect,” when the decision maker overly responds to the new information to estimate a post-test probability, without regard for the pretest probability. For example, troponin tests may be positive because of renal failure or sepsis in patients with a low pretest probability of acute thrombotic myocardial infarction. Taking the test result at face value and initiating therapy such as anti-thrombotic drug therapy in such a patient would be an example of base-rate neglect.

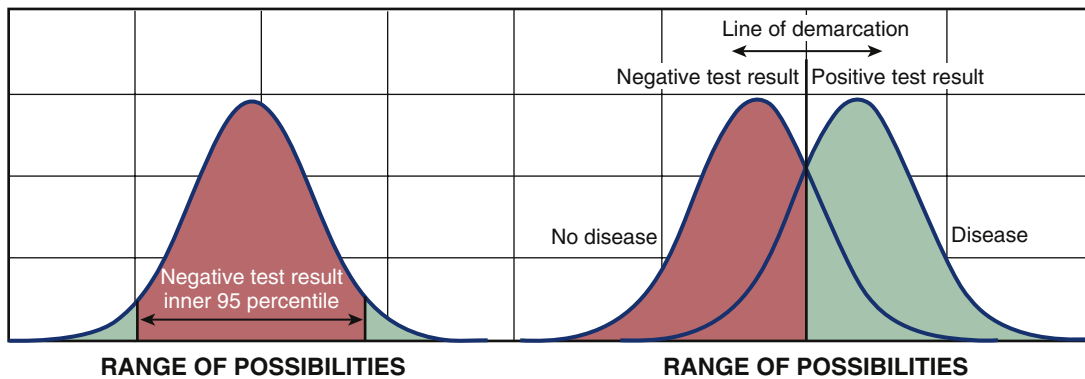


FIGURE 5.2 Left panel shows how the normal range of a test result is defined as the inner 95th percentile of a presumably normal population. Right panel shows how a normal and abnormal test result is defined by the line of demarcation between distributions of normal and abnormal test subjects, as defined by another independent “gold standard” test.

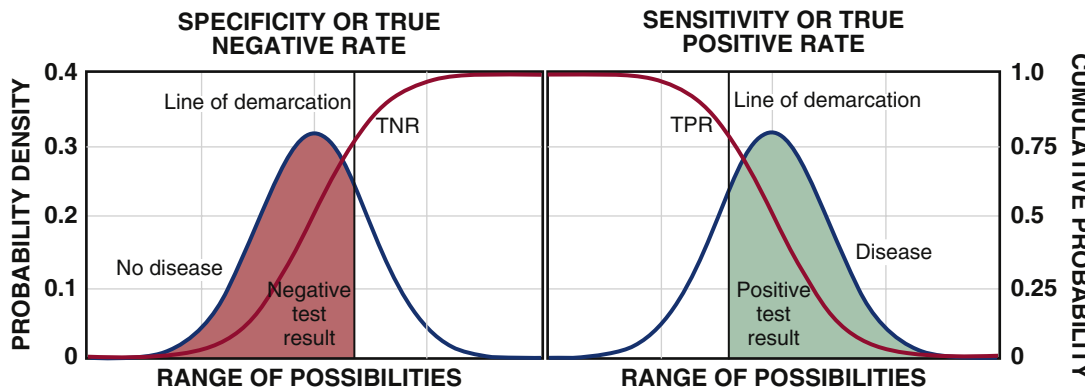


FIGURE 5.3 Distributions of normal and abnormal test subjects are shown separately to demonstrate that the true negative rate (TNR, or specificity) is the cumulative probability of a negative test result (red curve) in a distribution of subjects without disease (blue curve, left panel) and the true positive rate (TPR, or sensitivity) is the cumulative probability of a positive test result (red curve) in a distribution of subjects with disease (blue curve, right panel), depending on the location of the line of demarcation.

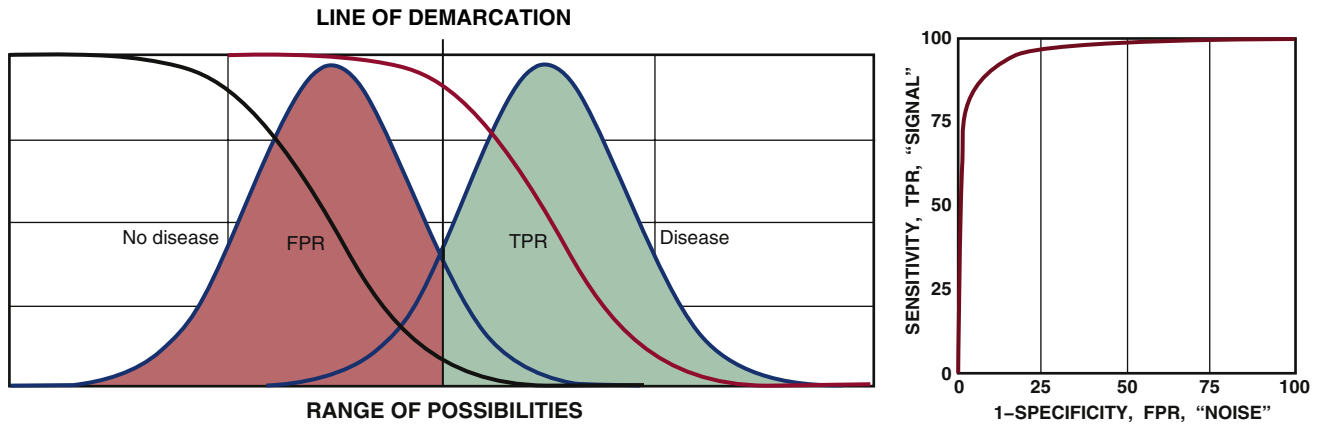


FIGURE 5.4 Distributions of patients without and with disease are shown in the left panel (blue curves) along with the true positive rate (TPR, or sensitivity, red curve) and the false-positive rate (FPR, or $1 - \text{specificity}$), black curve). The right panel shows the TPR (sensitivity or “the signal”) on the y-axis plotted against the FPR ($1 - \text{specificity}$) or “the noise”) on the x-axis, across the range of possibilities, depending on the location of the line of demarcation.

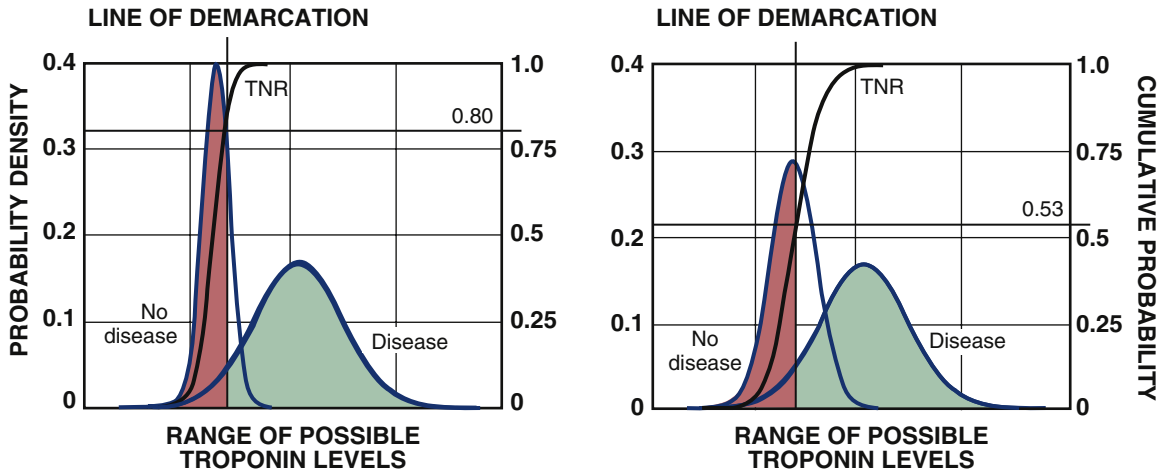


FIGURE 5.5 Distributions of patients without and with disease are shown by the blue curves. True negative test results are shown in red and true positive results are shown in green. The true negative rates (TNR, or specificity) are shown by the black cumulative probability curves. The left panel shows the results when the test is ordered on a narrowly defined population of subjects and the right panel shows the results when the test is ordered on a broadly defined population of test subjects, resulting in spectrum bias and a marked decrease in specificity of the test (80% to 53% in this example).

Likelihood ratios are useful for Bayesian reasoning. The advantage of likelihood ratios is that, unlike sensitivity and specificity, they are dimensionless numbers, so the need for keeping track of what is in the numerator and denominator is alleviated. Likelihood ratios give a measure of the persuasiveness of a positive and negative test result, and can be used intuitively, or used to actually calculate posttest odds.

A likelihood ratio is defined as the percentage of patients with a disease who have a given test result divided by the percentage of patients without disease who have that same test result. Thus, a positive likelihood ratio is the percentage of patients with disease with a positive test result divided by the percentage of patients without disease with a positive test result (TPR/FPR , or $\text{sensitivity}/[1 - \text{specificity}]$). A negative likelihood ratio is the percentage of patients with disease with a negative test result divided by the percentage of patients without disease with a negative test result (FNR/TNR , or $[1 - \text{sensitivity}]/\text{specificity}$). It is easy to calculate the positive and negative likelihood ratios from sensitivity and specificity. Once calculated, these numbers can be used to multiply the pretest odds to calculate the posttest odds of a diagnosis. They are multipliers, so a higher positive likelihood ratio, and a lower negative likelihood ratio (which is a fraction) have stronger multiplying effects. A likelihood ratio that is close to 1 is weak because it would have very little multiplying effect, meaning it has little effect on the pre-test assessment.

Figure 5.6 shows how the probability estimate of a diagnosis can shift depending on a test result. After choosing a pre-test probability estimate on the x-axis, one can trace up to either the upper curve for a positive test result or the lower curve for a negative test result, then trace over to the y-axis to read the post-test probability estimate. The diagonal line shows that there would be no change in probability for

a test with a likelihood ratio of 1. A higher positive likelihood ratio or a lower negative likelihood ratio would result in positive or negative test result curves with greater deviation from the diagonal line, representing a greater shift in the post-test probability estimate based on the test result.

Some tests are asymmetrical, meaning that either their positive or negative likelihood ratio is stronger. For example, Figure 5.7, Panel A shows the probability of congestive heart failure based on congestion on a chest x-ray, which has a very strong positive likelihood ratio of 13.5 and a relatively weak negative likelihood ratio of 0.48.³ This reflects the fact that the chest x-ray is highly specific but not very sensitive for heart failure. In other words, congestive findings on a chest x-ray are highly suggestive of heart failure, but their absence is not strong reassurance about the lack of heart failure. Tests that are highly specific are better for ruling in a diagnosis and this can be remembered using the mnemonic “SpPin.” (Highly specific tests, if positive, are good for ruling in.)

On the other hand, Figure 5.7, Panel B shows that a D-dimer for pulmonary embolus has a very strong negative likelihood ratio of 0.09 and a modest positive likelihood ratio of 1.7.³ This reflects the fact that a D-dimer is highly sensitive but not very specific for a pulmonary embolus. Tests that are highly sensitive are better for ruling out a diagnosis and this can be remembered using the mnemonic “SnNout.” (Highly sensitive tests, if negative, are good for ruling out.)

The likelihood ratios, however, are only as useful as the sensitivity and specificity that are used to calculate them. They give an approximate quantitative estimate of the strength of new information that provides a mechanism for calibrating intuitive probability estimates.

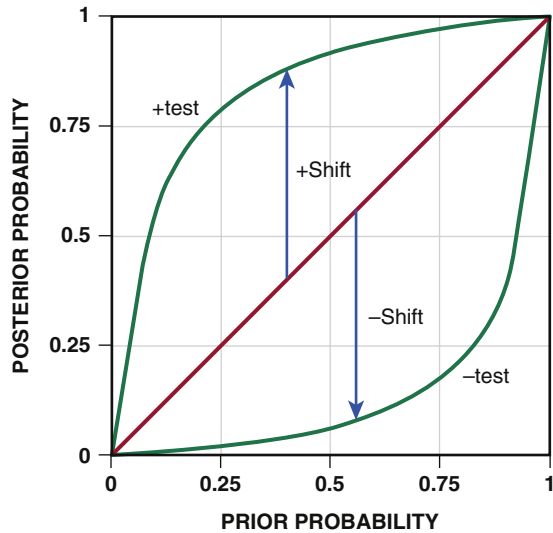


FIGURE 5.6 Leaf plot showing how the probability estimate of a diagnosis can shift depending on a test result. After choosing a pre-test probability estimate on the x-axis, one can trace up to either the upper curve for a positive test result or the lower curve for a negative test result, then trace over to the y-axis to read the post-test probability estimate.

TEST ORDERING STRATEGIES

Clinical reasoning should guide not only test interpretation, but also test ordering. Tests that are ordered for good reasons are more conclusive, and tests that are ordered indiscriminately can cause clinicians to make poor judgments. Ideally, a test should be used to validate or reject an articulated hypothesis—a plausible conjecture that is generated by a patient's condition. Ideally clinicians think ahead about what they would do with test results.

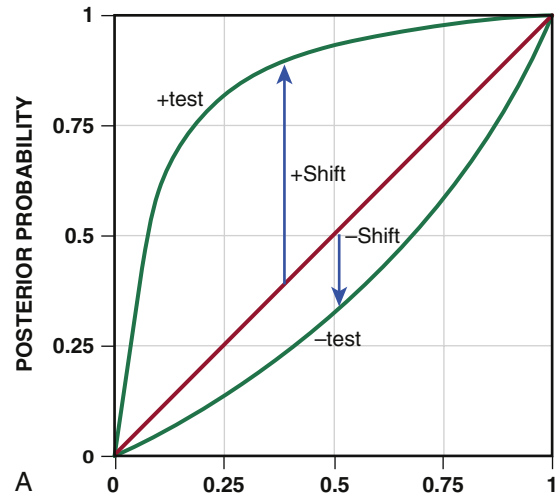
To aid with test selection and avoid over-testing, the American College of Cardiology (ACC) and other organizations have developed appropriate use criteria to guide clinicians' decisions about ordering selected cardiac tests.¹¹ This effort is driven by both a need to avoid excessive false-positive test results and also the need to contain the costs of medical care. The goal of appropriate use guidelines is to reduce overuse errors and to maximize the value of diagnostic testing and procedures. The general principle of any test-ordering strategy is that a plausible hypothesis (a provisional diagnosis) should be formulated first, followed by testing. The appropriate use criteria are designed to avoid testing when the results are unlikely to improve patient care and outcomes.

PREDICTING RISK

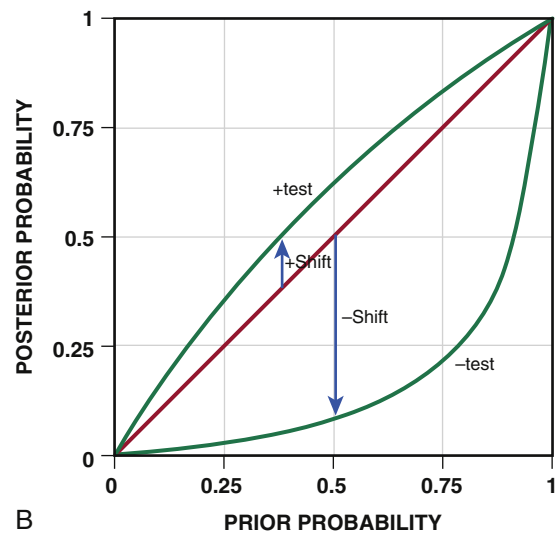
Recent ACC/American Heart Association (AHA) guidelines have promoted the provision of preventive treatments according to an individual's risk of adverse outcomes.¹² The premise is that low-risk people have little to gain by preventive interventions, while high-risk individuals may have a lot to gain. These guideline recommendations emphasize the need to consider categories based on estimates of risk and prognosis, rather than merely diagnostic labels, such as hyperlipidemia. It is important for clinicians to understand the provenance of the risk scores and their performance, including in diverse populations, to know whether the tools are useful. After calculating the risk, the challenge for clinicians is communicating risk to patients in an understandable fashion. Investigators have provided infographics that can communicate risk and risk reduction in order to facilitate a discussion regarding long-term treatment options to diminish risk, and to compare the degree of risk reduction with potential side effects and costs of treatment (see Fig. 5.8). Because clinicians vary in their use of qualitative terms such as "high risk," there is a need to provide clear and understandable quantitative estimates.¹³

THERAPEUTIC DECISIONS

A preventive or therapeutic decision is a structured choice. These decisions require medical knowledge and a balanced sense of risks and benefits, as well as knowledge of patients' preferences, to make optimal therapeutic decisions.



A




B


FIGURE 5.7 **A**, Probability estimate shift for a diagnostic test with a very strong positive likelihood ratio and a relatively weak negative likelihood ratio. **B**, Probability estimate shift for a diagnostic test with a very strong negative likelihood ratio and a modest positive likelihood ratio.

Clinical trials report the average risk of an outcome for patients in a treatment group and in a comparison group. There may be heterogeneity of the treatment effect, in which some patients may receive a marked benefit and others receive no benefit at all. Subgroup analysis and tests for interaction can provide hints, but usually heterogeneity of treatment effect is not readily apparent, creating a challenge for clinicians trying to personalize treatment decisions. In a key example of heterogeneity, fibrinolytic therapy was effective in the treatment of suspected AMI and subgroup analyses revealed the benefit to be substantial in patients with ST-elevation but not in those without it. The challenge is that subgroup analyses introduce the possibility that associations have occurred only by chance. In the Second International Study of Infarct Survival (ISIS-2), the authors provided perspective on subgroup analyses by demonstrating that patients born under the astrological signs of Gemini or Libra were significantly less likely to benefit from fibrinolytic therapy. Thus, subgroup analysis is capable of producing important insights, but must be interpreted with caution.¹⁴


A weakness of relative benefit estimates is that they do not convey information about what is achieved for patients at varying levels of risk. A small relative reduction in risk may be meaningful for a high-risk patient, while a large relative reduction may be inconsequential for a very low-risk patient. Absolute risk reduction, the difference between two rates, varies with the risk of an individual patient. For example, a risk ratio of 2.0 does not distinguish between baseline risks of 80% and



 Among untreated people, 80 people out of 100 will likely not develop an adverse cardiovascular outcome.

 20 people out of 100 will likely develop an adverse cardiovascular outcome.



 Among patients treated with a statin drug, 84 people out of 100 will likely not develop an adverse cardiovascular outcome.


 16 people out of 100 will likely develop an adverse cardiovascular outcome.

FIGURE 5.8 Infographic showing a format for demonstrating risk and benefit to patients. The infographic shows a baseline risk of a cardiovascular endpoint in patients with a baseline risk of 20%, and the 20% relative risk reduction with statin therapy.

40% and between 0.08% and 0.04%. In one case, the absolute difference is 50% (5000 per 10,000) and in the other, it is 0.05% (5 per 10,000). In one case, 1 person out of 2 benefits and in the other, 1 out of 2000 benefits. Unfortunately, absolute benefit is not emphasized adequately in many articles.¹⁵

Risk prediction is critically important for calculating the expected absolute risk reduction. In recent years, many tools have been developed to assist in the rapid assessment of patient risk, with variable uncertainty about their comparative performance.^{15,16}

In evaluating studies of risk prediction, it is important to consider whether the approach has been validated in populations similar to the patients to whom it is applied in practice. The predictors should be collected independently of knowledge of the outcome. The outcome and timeframe should be appropriate for clinical decisions and the value of the prediction should be clear. Appropriate risk prediction can assist in calculation of absolute benefit and put the balance of risks and benefits of an intervention in proper perspective.

Several studies have shown a risk-treatment paradox in which the higher-risk patients are least likely to receive interventions that are expected to provide a benefit. This pattern is paradoxical because the high-risk patients would be expected to have the most to gain from an intervention that reduces risk, assuming that the relative reduction in risk is constant across groups defined by their baseline risk. The source of the paradox is not known, although some have suggested that it is related to an aversion to the treatment of patients with a limited functional status, or concern for greater degree of harm from the same therapy.

Cardiovascular drugs and procedures are often double-edged swords, having both benefit and harm. Also, patients may have strong preferences about potential benefit and harm. For example, a patient may have a strong fear of a side effect such as a cerebrovascular accident that may overwhelm other considerations about a treatment decision. It is important to engage patients and families in a discussion to explain the considerations that go into therapeutic decisions, particularly for nuanced decisions about treatments that have substantial risks in addition to potential benefits.

Absolute risk reduction is better than relative risk reduction for estimating a treatment effect. The inverse of the absolute risk reduction, which is a term called number needed to treat (NNT), is even more intuitive.³

Consider a trial with a combined event rate of 10% in the treatment group and a 15% risk in the control group, giving an absolute risk reduction of 5%. This means that 5 events are avoided for every 100 patients in the treatment group. The reciprocal of this relationship indicates that there would be 100 patients treated for every 5 events avoided. By dividing 100 by 5, which reduces the denominator to 1, there would be 20 patients treated per 1 event avoided. Thus, the NNT is 20. For NNT, the smaller the number, the better.

NNT and absolute risk reduction depend on both the relative risk reduction and the baseline risk. For conditions with a high baseline risk, the NNT can become very small (desirable). As an extreme example, for a patient with ventricular fibrillation, the baseline risk of dying without defibrillation is 100%, making the NNT for defibrillation (if always effective) equal to 1.

Primary prevention with statin drugs has a relative risk reduction of about 20% over the several-year course of a typical prevention trial.¹⁷ The absolute risk reduction and NNT depend upon the baseline risk, which varies depending on a number of factors. At a baseline risk of 7.5%, the absolute risk reduction would be 1.5% and the NNT would be 67, a fairly high number, which suggests marginal benefit at this level of baseline risk.

NNT is a useful intuitive tool for comparing the efficacy of various treatment strategies. NNT is also a useful way to summarize the findings of a clinical trial in a single declarative sentence. For example, the PARTNER-3 trial had an NNT of 16, meaning one would need to treat 16 low-risk aortic stenosis patients with a transcatheter aortic valve replacement to prevent one composite endpoint of death, stroke, or rehospitalization over 1 year.¹⁸ The EMPEROR-Reduced trial had an NNT of 19, meaning one would need to treat 19 patients with class II to IV heart failure and an ejection fraction of $\leq 40\%$ with a sodium-glucose cotransporter 2 (SGLT2) inhibitor for 16 months to prevent one death or hospitalization for worsening heart failure.¹⁹ With NNT, a single sentence can provide the trial name, the magnitude of the treatment effect, the trial's entry criteria, the study drug or intervention, the duration of the trial, and the outcome measure. NNTs of 16 and 19 suggest that these treatments are strongly recommended, although some patients may not consider it worth going through the treatment if 15 of 16 people have the same outcome regardless of whether they received the intervention.

NNT is also a very personal notion of the probability of a treatment effect. Imagine bringing 19 untreated patients with congestive heart failure and an ejection fraction of $\leq 40\%$ into a room and saying, "If all of you start on an SGLT2 inhibitor, over the next 16 months, one of you will experience the benefit." Capturing the essence of a treatment effect with NNT is a useful way to intuitively convey the impact of a treatment effect. This knowledge, packaged in a way that is more intuitive, can make it easier to combine this medical knowledge with the preferences and values of individual patients to make the best therapeutic decisions.

Nevertheless, there are limitations to NNT. NNT is an index of an average treatment effect over time and does not provide information about whether the treatment effect is immediate, delayed, or highly variable. NNT does not provide information about whether there is meaningful heterogeneity in effect among different subgroups, as the NNT often is calculated based on an assumption of a uniform effect of the therapy, with the NNT just varying based on the baseline risk.

DECIDING WHEN TO CHANGE CLINICAL PRACTICE BASED ON NEW CLINICAL RESEARCH FINDINGS

Science is a quantitative discipline that uses numbers to measure, analyze, and explain nature. Evidence-based medicine has been defined by David Sackett as “the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients.” To practice evidence-based medicine, clinicians must remain vigilant, constantly monitoring for new research findings, accompanied by a basic knowledge of statistics to make proper inferences from clinical research.

When using statistics to compare two groups, the standard method is to assume that there is no difference between the two groups, the so-called null hypothesis. The trial results are reported, along with a p value, which is the probability of deriving the difference reported in the trial, or a more extreme difference, given the assumption that the null hypothesis is true (i.e., there is no real difference between the groups). When a trial is designed, the investigators estimate the sample sizes that are required to avoid claiming that there is a difference between treatment groups when there really is no difference (a type I error or alpha error) or claiming that there is no differences between treatment groups when there really is one (a type II error or beta error). Similar to a clinical test like a stress test that can have false-positive and false-negative test results, clinical trials can have false-positive results (alpha errors) and false-negative results (beta errors). A trial with adequate sample size and rigorous statistical methods should allow investigators to avoid these errors.

When a trial is designed, the alpha level is usually set at 0.05. If the p value of the observed data is less than 0.05, one can conclude that a very improbable event occurred, a less than 1-in-20 event, assuming the null hypothesis is valid. According to the frequentist notion of statistics, one imagines that repeating a trial many times would create a distribution of possible trial results. The p value tells us where the observed results of a particular trial would sit in that imaginary distribution of trial results.

Because the p value is so commonly used in clinical research, clinicians need to be aware of several key issues. First, the threshold of 0.05 for statistical significance is arbitrary. A p value of 0.04 implies that the data could occur 4% of the time if the null hypothesis is true and a p value of 0.06 would suggest the data would occur 6% of the time. Is the difference between 6% and 4% enough to reject the null hypothesis in one case and accept it in another? Clinicians should understand that p values are continuous values and are just one piece of information needed to assess a trial. Second, p values do not inform clinical importance. A large study sample can produce a small p value despite a clinically inconsequential difference between groups. Clinicians need to examine the size of the effects in addition to the statistical tests of whether the results could have occurred by chance.

ENDPOINTS

In evaluating evidence, clinicians should be particularly attuned to the outcomes that are assessed. Ideally, interventions are assessed for their effect on a patient’s quality or quantity of life. Many studies use surrogate outcomes, measures that are more distally related to the patient experience but might be related to the likelihood that a patient’s quality or quantity of life will be affected. These surrogate outcomes often reflect information about a patient’s biology, and in epidemiologic studies, these outcomes may have prognostic value. However, it is not possible to know that an intervention that modifies a surrogate outcome has the expected effect on patients. There are many examples in medicine of changes in surrogate measures that did not translate into benefits for patients.

NONINFERIORITY TRIALS

Most randomized trials are designed to show the superiority of a treatment over placebo. However, some conditions already have treatments with proven benefit, making it unethical to design a trial that compares a new treatment with placebo. For example, for chronic atrial fibrillation, it was not possible to test newer oral anticoagulant drugs against a placebo arm that would have withheld the proven benefit of warfarin.

For these situations, investigators use a noninferiority trial. The premise is to show that a given treatment is at least no worse than the standard of care by more than a predefined investigator-selected margin (the treatment could be slightly worse, or even be superior for efficacy). However, because the new treatment has other ancillary advantages (e.g., fewer side effects, better costs or tolerability), it could become a reasonable alternative to the previous standard of care. This trial design requires making assumptions about the margin of decreased efficacy that would be considered acceptable before considering using a new treatment rather than an established treatment with known efficacy. Noninferiority trials are also subject to several other biases that are not seen with typical superiority trials.²⁰

OBSERVATIONAL RESEARCH

There are other situations in which a randomized controlled trial is impossible, and observational studies such as case-control studies or longitudinal cohort studies are required. Randomized controlled trials have the advantage of a controlled experiment that eliminates potential biases but have the limitation of narrowly defining a study population, which may affect generalizability. Observational studies have the advantage of observing large groups of unselected subjects in the real-world setting but have the disadvantage of potentially unrecognized and unmeasured sources of bias that could produce misleading study results. There is great heterogeneity in the quality of the data and methods used in observational research and there is a need to be discerning when evaluating observational research studies.^{21,22}

SHARED DECISION-MAKING

The principle of autonomy maintains that patients retain control over their bodies and must consent to undergo interventions, except in rare circumstances. Informed consent is the cornerstone of this concept.²³ Unfortunately, there is little consensus about how best to involve patients actively in decision-making. Nevertheless, given the need to align goals of therapy with the patient’s preferences and values, it is important to engage them as effectively as possible. This approach is most appropriate for major decisions, those with intermediate or low certainty, and those that are not emergent.

Informed consent is a critical part of medicine, but one that has many gaps. Documentation of informed consent suggests that physicians often do not communicate in ways that address key topics necessary for informed choice.²⁴ It has been suggested, for example, that informed consent should be obtained with enough time before elective procedures for people to contemplate the decisions.²⁵ The consent about procedures should contain an easy-to-understand description of the procedure, potential benefits, potential risks, alternative strategies, out-of-pocket costs, and the experience of the health care team in performing the procedure.²⁶

There are many aspects of communicating risks and benefits. First, this information takes many forms. The dimensions of risk and benefit include their identity, permanence, timing, probability, and value to an individual patient. All should be considered in decision-making. Unfortunately, there is relatively little evidence to guide physicians about how best to convey risks to patients.

It is known that patients do not always understand benefit and risk well. Many studies of patients undergoing revascularization procedures have shown that the patients can make assumptions about the survival benefits of cardiac procedures and patients often have a poor understanding of the potential for complications. There have been calls to improve medical education and medical reporting to address this problem of “risk illiteracy.”²⁷

The manner in which information is presented may influence patients. Like physicians, patients are also susceptible to framing effects. Patients tend to be more likely to choose a therapy that is presented as having an advantage over an alternative in relative rather than absolute terms. The relative effect is almost always much greater than the absolute change. Patients may also be influenced by the order in which information is provided.

Some techniques have been proposed to help clinicians convey risk.²⁸ First, clinicians should avoid descriptive terms only because they may not have a consistent meaning to patients. Terms such as “low risk” may be difficult for people to interpret. If clinicians express



risk as ratios, they should use a consistent denominator (e.g., 40 out of 1000 and 5 out of 1000 instead of 1 in 25 and 1 in 200). Clinicians should offer multiple perspectives, revealing multiple ways of thinking about risk. They should use absolute numbers and natural frequencies (e.g., 1 out of 20), not relative risks or percentages. Visual aids are useful, if available, as poor numeracy or literacy skills may be a barrier for many patients. Many patients do not understand risk communication formats. In addition, clinicians should recognize that information and data are not the same, and it is incumbent on the clinician to communicate health information that is meaningful to the patient.

Figure 5.8 shows a format for demonstrating risk and benefit to patients. The infographic shows a baseline risk of a cardiovascular endpoint in patients with a baseline risk of 20%, and the 20% relative risk reduction with statin therapy.¹⁷

Shared decision-making can be understood as having five phases: assess, advise, agree, assist, and arrange. First, the clinician must assess the patient. Then, the clinician should advise the patient of the options, with their benefits and risks. Next, the clinician and patient should agree on a plan that is aligned with the patient's preferences and values. The clinician should then assist the patient in implementing the plan. Finally, the patient and clinician arrange follow-up.^{29,30}

Clinical decision-making also involves other members of a cardiovascular team.^{31,32} The model of team-based care has now become standard practice, bringing Advanced Practice Practitioners (nurse practitioners, physician assistants, and clinical pharmacists) into the decision-making process for cardiovascular care. Good communication and coordination, and the mantra of "shared goals and clear roles" are important to assure optimal team-based shared decision-making.

MONITORING THE QUALITY OF CLINICAL DECISIONS

Getting the right care to the right patient at the right time every time requires good judgment. Learning the basic competencies of good judgment and step-by-step methods of clinical reasoning can help practitioners monitor the quality of their decisions. Knowledge about clinical reasoning is a structural attribute that can lead to more reliable processes and better clinical outcomes. Awareness of the logic, probability theory, and cognitive psychology of clinical reasoning can provide a theoretical foundation for better clinical practice.

Cognitive science provides justification for many of the good habits that are part of practice, such as consistently performing a standardized history and physical examination and a conscious habit of listing a differential diagnosis. Cognitive psychologists emphasize that measurement and feedback is a crucial process for the development of expert intuition, which is so often necessary for clinical decisions.

System 1 and System 2 Thinking

Cognitive psychologists describe two general thinking modes that people use to make decisions: System 1 and System 2 thinking.⁶ System 1 thinking is highly intuitive and fast, but prone to jumping to conclusions. System 2 thinking is analytical and logical, but slow, effortful, and has difficulty with uncertainty. Used together, System 2 thinking provides a double check for System 1 thinking, and System 1 thinking provides a work-around when System 2 thinking is constrained by uncertainty. Cardiology decisions require both thinking modes, and expert clinicians are able to use a balance of intuition and analytical thinking to make optimal decisions. Calibrating intuitive thinking and organizing thinking by thoughtfully monitoring clinical decisions (so called "meta-cognition") are key to good clinical practice.

Some psychologists describe three general types of fallacies: (1) hasty judgments, (2) biased judgments, and (3) distorted probability estimates.⁶ Hasty judgments occur when System 1 thinking is unmonitored. For example, premature closure of a diagnostic exercise, without the use of a differential diagnosis, or becoming anchored on a diagnosis can lead to a misdiagnosis. Biased judgments occur when

unconscious thoughts influence ideas, emotions, and actions. This can take the form of priming, stereotyping, overconfidence, risk aversion, or dread. Emotions can have a halo effect, influencing clinical thinking in imperceptible ways. Exaggerated fear of malpractice, financial incentives, and conflict of interest can adversely affect decisions. Decision-makers tend to overweigh the probabilities of events or propositions at the extremes. At one end, clinicians can develop an illusion of certainty, which creates certainty about something that, objectively, is not certain at all. At the other extreme is the possibility effect, which creates the impression that highly improbable events or propositions are quite probable. Knowledge of these fallacies can help clinicians develop habits that will improve the quality of their clinical reasoning.

A recent Institute of Medicine report brought attention to the high prevalence of diagnostic error.³³ Renewed interest has centered around defining diagnostic errors and measuring error rates.

A number of authors have suggested that diagnostic errors are due to cognitive biases.^{33,34} To avoid cognitive biases, the cognitive psychologist Kahneman provides the following advice: "The way to block errors that originate in System 1 is simple in principle: recognize the signs that you are in a cognitive minefield and slow down, and ask for reinforcement from System 2."³⁶ Evans, another cognitive psychologist, counters this advice by stating, "perhaps the most persistent fallacy in the perception of dual-process theories is the idea that Type 1 processes (intuitive, heuristic) are responsible for all bad thinking and that Type 2 processes (reflective, analytic) necessarily lead to correct responses... So ingrained is this good-bad thinking idea that some dual process theories have built it into their core terminology."³⁵

In fact, a number of studies show that faster response times often lead to a correct diagnosis. One study showed that when test subjects were instructed to go slow and be more analytic, there was no effect on the accuracy of diagnosis.³⁶

Some investigators have advocated teaching physicians a large number of potential biases and encouraging them to routinely "de-bias" their thinking in practice.³⁴ However, studies that examined educational efforts to teach cognitive biases to medical trainees found that the instruction had no effect.^{37,38} One study showed that physicians were unable to agree on what bias contributed to a diagnostic error and the study subjects were themselves affected by hindsight bias.³⁹

Norman and others have argued that diagnostic error is commonly due to knowledge deficits or inability to mobilize necessary knowledge rather than flawed reasoning processes and biases.³⁸ Using deliberate reflection to help mobilize the necessary knowledge, as well as the use of checklists have been shown to be beneficial for reducing diagnostic error.⁴⁰ Improving diagnostic reasoning is an important goal and more research is needed to determine the most effective strategies for improvement.

Monitoring the quality of therapeutic decision-making is a broad topic and has been the focus of quality improvement efforts for decades. Explicit therapeutic decisions produce overt actions, and therefore can be easily monitored. Implicit bias, however, can have subtle effects on clinical decision-making, resulting in disparities in diagnosis, treatment, and outcomes. Implicit bias occurs when one's judgment is affected by unconscious associations between a class such as race and learned cultural prejudices and stereotypes.⁴¹ Implicit bias related to race, sex, age, obesity, and other patient factors can unconsciously bias medical decisions. Conscious awareness of the problem through cultural competency training is helpful but may not be sufficient to alleviate implicit bias. Investigators have suggested that articulating egalitarian goals, identifying common identities with the patient, counter-stereotyping, and trying to see things from the patient's perspective are conscious mental strategies that can diminish implicit bias.⁴²

TEACHING CLINICAL REASONING

Clinical reasoning can be learned through experience and continuously improved through deliberate practice and reflection. Although clinical reasoning can be learned, there are remaining questions about whether clinical reasoning can be formally taught through explicit instruction. Most teaching of clinical reasoning is through experience

in the setting of clinical rotations. Formal teaching of reasoning has been promoted for centuries, but Richard Nesbit and others have posited that teaching abstract rules of reasoning fell into disfavor in the 20th century (see Classic References). The prevailing notion was that learners do not use abstract reasoning rules, but rather use domain-specific empirical rules that are better learned by experience, not instruction. There is evidence, however, that statistical heuristics, pragmatic inferential rules, and metacognitive strategies can be taught even with brief formal training.⁴³

A recent study examined whether Bayesian reasoning can be taught.⁴⁴ The study showed a modest advantage for medical students randomized to receive theoretical instruction on concepts of Bayesian reasoning, suggesting that a conceptual framework can modestly improve clinical decision-making.⁴⁵

Clinical reasoning is not a specific problem-solving skill that easily transfers from problem to problem, but rather is dependent on learned habits and a conceptual framework that enable the clinician to access, organize, and use both experiential knowledge and formal knowledge to address a clinical problem. Interleaving instruction on concepts of reasoning throughout training may strengthen the associations and provide a scaffolding for experiential knowledge to help learners make the most of their experience, calibrate their intuitive judgments, and become better at clinical reasoning.

CONCLUSION

Reliable decision-making is fundamentally important for high-quality, patient-centered medical care. Clinical decisions in cardiology are often time-sensitive, complex, and uncertain, demanding good clinical reasoning. Logic, probability theory, and cognitive science can provide a framework for good clinical reasoning. The ability to read, understand, and critique the literature is also essential. Knowledge of the components of clinical reasoning is crucial for clinical practice, for team-based care, and for shared decision-making. The ability to reason and the ability to use reasoning to stay current and monitor one's performance are the essence of professionalism. Ensuring that the patient is a participant and is part of all decisions is also essential. Integrating scientific knowledge and calibrated intuition with a patient's personal preferences and values can provide the highest quality of care.

CLASSIC REFERENCES

- Simon H. Invariants of human behavior. *Annu Rev Psychol.* 1990;41:1–20.
 Montgomery K. *How Doctors Think: Clinical Judgment and the Practice of Medicine.* Oxford, New York: Oxford University Press; 2006.
 Elstein AS, Shulman LS, Sprafka SA. *Medical Problem Solving: An Analysis of Clinical Reasoning.* Cambridge, MA: Harvard University Press; 1978.
 Diamond GA, Forrester JS. Analysis of probability as an aid in the clinical diagnosis of coronary artery disease. *N Engl J Med.* 1979;300:1350–1358.
 Nisbett RE, Fong GT, Lehman DR, Cheng PW. Teaching reasoning. *Science.* 1987;238:625–631.

REFERENCES

Access supplemental references online at Elsevier eBooks for Practicing Clinicians.

- Kassirer JP, Wong JB, Kopelman RI. *Learning Clinical Reasoning.* Baltimore, MD: Lippincott Williams & Wilkins Health; 2010.
- Sox HC, Higgins MC, Owens DK. *Medical Decision-Making.* 2nd ed. West Sussex, UK: Wiley-Blackwell; 2013.
- Brush JE. *The Science of the Art of Medicine: A Guide to Medical Reasoning.* Manakin-Sabot, VA: Dementi Milestone Publishing, Inc.; 2015.
- Norman GR, Eva KW. Diagnostic error and clinical reasoning. *Med Educ.* 2010;44:94–100.
- Dharmarajan K, Strait KM, Tinetti ME, et al. Treatment for multiple acute cardiopulmonary conditions in older adults hospitalized with pneumonia, chronic obstructive pulmonary disease, or heart failure. *J Am Geriatr Soc.* 2016;64:1574–1582.
- Kahneman D. *Thinking, Fast and Slow.* New York: Farrar, Straus and Giroux; 2011.
- Gigerenzer G, Gaissmaier W. Heuristic decision making. *Ann Rev Psychol.* 2011;62:451–482.
- Brush Jr JE, Sherbino J, Norman GR. How expert clinicians intuitively recognize a medical diagnosis. *Am J Med.* 2017;130:629–634.
- Brush Jr JE, Krumholz HM, Greene EJ, Dreyer RP. Sex differences in symptom phenotypes among patients with acute myocardial infarction. *Circ Cardiovasc Qual Outcomes.* 2020;13:e005948.
- Manrai AK, Funke BH, Rehm HL, et al. Genetic misdiagnoses and the potential for health disparities. *N Engl J Med.* 2016;375:655–665.
- Hendel RC, Lindsay BD, Allen JM, et al. ACC appropriate use criteria methodology: 2018 update: a report of the American College of Cardiology Appropriate use criteria task force. *J Am Coll Cardiol.* 2018;71:935–948.
- Goff Jr DC, Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association task force on practice guidelines. *J Am Coll Cardiol.* 2014;63:2935–2959.
- Spiegelhalter DJ. *The Art of Statistics: How to Learn from Data.* First US edition. New York: Basic Books; 2019.
- O'Gara PT, Kushner FG, Ascheim DD, et al. 2013 ACCF/AHA guideline for the management of STElevation myocardial infarction: executive summary: a report of the American College of Cardiology Foundation/American Heart Association task force on practice guidelines. *J Am Coll Cardiol.* 2013;61:485–510.
- Damen JA, Hooft L, Schuit E, et al. Prediction models for cardiovascular disease risk in the general population: systematic review. *BMJ.* 2016;353:i2416.
- McNamara RL, Kennedy KF, Cohen DJ, et al. Predicting in-hospital mortality in patients with acute myocardial infarction. *J Am Coll Cardiol.* 2016;68:626–635.
- Stone NJ, Robinson JG, Lichtenstein AH, et al. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol.* 2014;63:2889–2834.
- Mack MJ, Leon MB, Thourani VH, et al. Transcatheter aortic-valve replacement with a balloon-expandable valve in low-risk patients. *N Engl J Med.* 2019;380:1695–1705.
- Packer M, Anker SD, Butler J, et al. Cardiovascular and renal outcomes with empagliflozin in heart failure. *N Engl J Med.* 2020;383:1413–1424.
- Bikdeli B, Welsh JW, Akram Y, et al. Noninferiority designed cardiovascular trials in highest-impact journals. *Circulation.* 2019;140:379–389.
- Califf RM, Hernandez AF, Landray M. Weighing the benefits and risks of proliferating observational treatment assessments: observational cacophony, randomized harmony. *J Am Med Assoc.* 2020. <https://doi.org/10.1001/jama.2020.13319>.
- Schuemie MJ, Ryan PB, Pratt N, et al. Principles of Large-scale Evidence Generation and Evaluation across a Network Of Databases (LEGEND). *J Am Med Inform Assoc.* 2020;27:1331–1337.
- Spatz ES, Krumholz HM, Moulton BW. The new era of informed consent: getting to a reasonable-patient standard through shared decision making. *J Am Med Assoc.* 2016;315:2063–2064.
- Spatz ES, Bao H, Herrin J, et al. Quality of informed consent documents among US hospitals: a cross-sectional study. *BMJ Open.* 2020;10:e033299.
- Spatz ES, Suter LG, George E, et al. An instrument for assessing the quality of informed consent documents for elective procedures: development and testing. *BMJ Open.* 2020;10:e033297.
- Krumholz HM. Informed consent to promote patient-centered care. *J Am Med Assoc.* 2010;303:1190–1191.
- Gigerenzer G, Gray JAM. *Better Doctors, Better Patients, Better Decisions: Envisioning Health Care 2020.* Cambridge, MA: MIT Press; 2011.
- Navar AM, Wang TY, Mi X, et al. Influence of cardiovascular risk communication tools and presentation formats on patient perceptions and preferences. *JAMA Cardiol.* 2018;3:1192–1199.
- Krumholz HM. Variations in health care, patient preferences, and high-quality decision making. *J Am Med Assoc.* 2013;310:151–152.
- Krumholz HM, Barreto-Filho JA, Jones PG, et al. Decision-making preferences among patients with an acute myocardial infarction. *JAMA Intern Med.* 2013;173:1252–1257.
- Brush Jr JE, Handberg EM, Biga C, et al. 2015 ACC health policy statement on cardiovascular team-based care and the role of advanced practice providers. *J Am Coll Cardiol.* 2015;65:2118–2136.
- Rodgers GP, Linderbaum JA, Pearson DD, et al. 2020 ACC clinical competencies for nurse practitioners and physician assistants in adult cardiovascular medicine: a report of the ACC Competency Management Committee. *J Am Coll Cardiol.* 2020;75:2483–2517.
- National Academies of Sciences, Engineering, and Medicine. *Improving Diagnosis in Health Care.* Washington, DC: The National Academies Press; 2015. <https://doi.org/10.17226/21794>.
- Croskerry P. From mindless to mindful practice—cognitive bias and clinical decision making. *N Engl J Med.* 2013;368:2445–2448.
- Evans JS. Dual-processing accounts of reasoning, judgment, and social cognition. *Annu Rev Psychol.* 2008;59:255–278.
- Sherbino J, Dore KL, Wood TJ, et al. The relationship between response time and diagnostic accuracy. *Acad Med.* 2012;87:785–791.
- Sherbino J, Kulasegaram K, Howey E, Norman G. Ineffectiveness of cognitive forcing strategies to reduce biases in diagnostic reasoning: a controlled trial. *CJEM.* 2014;16:34–40.
- Norman GR, Monteiro SD, Sherbino J, et al. The causes of errors in clinical reasoning: cognitive biases, knowledge deficits, and dual process thinking. *Acad Med.* 2017;92:23–30.
- Zwaan L, Monteiro S, Sherbino J, et al. Is bias in the eye of the beholder? A vignette study to assess recognition of cognitive biases in clinical case workshops. *BMJ Qual Saf.* 2017;26:104–110.
- Schmidt HG, Mamede S. How to improve the teaching of clinical reasoning: a narrative review and a proposal. *Med Educ.* 2015;49:961–973.
- FitzGerald C, Hurst S. Implicit bias in healthcare professionals: a systematic review. *BMC Med Ethics.* 2017;18:19.
- Stone J, Moskowitz GB. Non-conscious bias in medical decision making: what can be done to reduce it? *Med Educ.* 2011;45:768–776.
- Nesbitt RE. *Mindware: Tools for Smart Thinking.* New York, NY: Farrar, Straus and Giroux; 2016.
- Brush Jr JE, Lee M, Sherbino J, et al. Effect of teaching Bayesian methods using learning by concept vs learning by example on medical students' ability to estimate probability of a diagnosis: a randomized clinical trial. *JAMA Netw Open.* 2019;2:e1918023.
- Kulasegaram KM, Chaudhary Z, Woods N, et al. Contexts, concepts and cognition: principles for the transfer of basic science knowledge. *Med Educ.* 2017;51:184–195.

Diagnostic Decisions

- Brush Jr JE, Sherbino J, Norman GR. How expert clinicians intuitively recognize a medical diagnosis. *Am J Med.* 2017;130:629–634.

ONLINE-SUPPLEMENTAL REFERENCES

46. Norman G, Eva K, Brooks L, Hamstra S. Expertise in medicine and surgery. In: Ericsson KA, et al., ed. *The Cambridge Handbook of Expertise and Expert Performance*. Cambridge; New York: Cambridge University Press; 2006.
47. Ericsson KA. An expert-performance perspective of research on medical expertise: the study of clinical performance. *Med Educ*. 2007;41:1124–1130.
48. Bowen JL. Educational strategies to promote clinical diagnostic reasoning. *N Engl J Med*. 2006;355:2217–2225.
49. Hacking I. *An Introduction to Probability and Inductive Logic*. Cambridge, U.K.; New York: Cambridge University Press; 2001.
50. Sackett DL, Haynes RB, Guyatt GH, Tugwell P. *Clinical Epidemiology: A Basic Science for Clinical Medicine*. 2nd ed. Boston: Little, Brown, & Co; 1991.
51. Simel DL, Rennie D, Keitz SA. *The Rational Clinical Examination: Evidence-Based Clinical Diagnosis*. New York: McGraw-Hill; 2009.
52. Sackett DL, Rosenberg WM, Gray JA, et al. Evidence-based medicine: what it is and what it isn't. *BMJ*. 1996;312:71.
53. Goodman SN. Toward evidence-based medical statistics. 1: the p value fallacy. *Ann Intern Med*. 1999;130:995–1004.
54. Bogardus Jr ST, Holmboe E, Jekel JF. Perils, pitfalls, and possibilities in talking about medical risk. *J Am Med Assoc*. 1999;281:1037–1044.
55. Epstein RM, Alper BS, Quill TE. Communicating evidence for participatory decision making. *J Am Med Assoc*. 2004;291:2359–2366.
56. Holmboe ES, Fiellin DA, Cusanelli E, et al. Perceptions of benefit and risk of patients undergoing first-time elective percutaneous coronary revascularization. *J Gen Intern Med*. 2000;15:632–637.
57. Whittle J, Conigliaro J, Good CB, et al. Understanding of the benefits of coronary revascularization procedures among patients who are offered such procedures. *Am Heart J*. 2007;154:662–668.
58. Malenka DF, Baron JA, Johansen S, et al. The framing effect of relative and absolute risk. *J Gen Intern Med*. 1993;8:543–548.
59. Sepucha KR, Levin CA, Uzogara EE, et al. Developing instruments to measure the quality of decisions: early results for a set of symptom-driven decisions. *Patient Educ Couns*. 2008;73:504–510.