Diagnostic reasoning in cardiovascular medicine

John E Brush Jr,[1,2] Jonathan Sherbino,[3,4] Geoffrey R Norman[3]

[1]Sentara Health Research Center, Norfolk, VA, USA
[2]Eastern Virginia Medical School, Norfolk, VA, USA
[3]McMaster Education Research, Innovation and Theory (MERIT) Program, McMaster University, Hamilton, ON, Canada
[4]Department of Medicine, McMaster University, Hamilton, ON, Canada
Correspondence to: J E Brush jebrush@sentara.com

Cite this as: *BMJ* 2022;376:e064389
http://dx.doi.org/10.1136/bmj-2021-064389

Series explanation: State of the Art Reviews are commissioned on the basis of their relevance to academics and specialists in the US and internationally. For this reason they are written predominantly by US authors.

ABSTRACT
Research in cognitive psychology shows that expert clinicians make a medical diagnosis through a two step process of hypothesis generation and hypothesis testing. Experts generate a list of possible diagnoses quickly and intuitively, drawing on previous experience. Experts remember specific examples of various disease categories as exemplars, which enables rapid access to diagnostic possibilities and gives them an intuitive sense of the base rates of various diagnoses. After generating diagnostic hypotheses, clinicians then test the hypotheses and subjectively estimate the probability of each diagnostic possibility by using a heuristic called anchoring and adjusting. Although both novices and experts use this two step diagnostic process, experts distinguish themselves as better diagnosticians through their ability to mobilize experiential knowledge in a manner that is content specific. Experience is clearly the best teacher, but some educational strategies have been shown to modestly improve diagnostic accuracy. Increased knowledge about the cognitive psychology of the diagnostic process and the pitfalls inherent in the process may inform clinical teachers and help learners and clinicians to improve the accuracy of diagnostic reasoning. This article reviews the literature on the cognitive psychology of diagnostic reasoning in the context of cardiovascular disease.

Introduction
An accurate and timely diagnosis is of paramount importance for patients with heart disease. A missed cardiac diagnosis can lead to harm to the patient, as well as dissatisfaction and life threatening ramifications.[1-3] Cardiac diseases are common, and cardiac diagnostic encounters are frequent, particularly in emergency departments where a missed cardiac diagnosis is a leading cause of malpractice litigation.[4 5] One cardiac diagnostic challenge is an ST elevation myocardial infarction (STEMI), for which a correct and timely diagnosis is critical for triggering emergency life saving interventions.[6 7] The door-to-balloon time for STEMI is a widely reported measure of a hospital's performance that is critically dependent on reliable diagnostic competence.[8] Diagnostic accuracy and timing are critical for cardiovascular patients, yet the process of diagnostic reasoning is underemphasized in cardiology training and continuing medical education, as it is in many specialty areas.[9]

Insufficient attention has been paid to the important area of diagnostic reasoning, which has been eclipsed in the cardiology literature by reports of large clinical trials and novel technological advances. Also, very few studies have examined how cardiovascular diagnostic strategies affect decisions about patients' management and outcomes, and costs, even though cardiovascular diagnostic tests and imaging are frequently used, have substantial impact, and are exceedingly costly.[10 11]

Medical error was brought to public attention in the United States and elsewhere more than two decades ago by the Institutes of Medicine's publication *To Err Is Human*.[12] A more recent Institutes of Medicine publication, *Improving Diagnosis in Health Care*, brought attention to the problem of diagnostic error with calls for better education about the diagnostic process, better measurement of diagnostic error, and more research and emphasis on diagnostic competency.[13]

Hiding in plain view has been the abundant literature in cognitive science, which has yielded important evidence on how clinical experts make a diagnosis. By "expert," we mean a clinician who has completed specialty training and is in practice, and so can be assumed to have the necessary knowledge and experience in their specialty; a "novice" is a learner in the early stages of training who has not yet acquired the necessary knowledge and experience for independent practice. Much of this literature has been published in education and psychology journals and may have escaped the attention of practicing cardiologists and clinical teachers. This review summarizes the evidence base about diagnostic reasoning in the context of cardiovascular disease, with the intention that increased awareness among clinicians and teachers will improve the quality of cardiovascular diagnostic reasoning.

Sources and selection criteria

This narrative review synthesizes the diagnostic reasoning literature from medicine and applies this evidence base to the domain of cardiovascular disease. Building on our knowledge of this literature, we did a broad, systematic search using the term "diagnostic reasoning" in Google Scholar and PubMed in all date ranges. We supplemented this search strategy by a hand search of the references of key articles. We achieved inclusion of identified articles by using an informal consensus approach based on an assessment of a study's impact and its methods, with preference given to experimental studies. We achieved organization and framing of key themes that emerged during the synthesis in an iterative fashion, by consensus. We gave precedence to references that support our shared view that the effectiveness of educational strategies and corrective measures for diagnostic reasoning should be subjected to formal evaluation.

Incidence of diagnostic error

The incidence of diagnostic errors can be estimated from several sources, including autopsy studies, surveys of patients, audits of diagnostic testing, and reviews of closed malpractice claims.[14] One retrospective analysis of internal medicine cases estimated that the rate of diagnostic error was very high, possibly in the 10-15% range.[15] This and other reports have noted that estimating rates of diagnostic error on the basis of retrospective review has limitations owing to detection and reporting biases.[16] Another observational study identified missed diagnoses of acute myocardial infarction (AMI) by counting the number of patients who returned to an emergency department, which would likely undercount the number of missed diagnoses.[3] Another limitation is that, over time, diseases can progress and diagnostic evidence can accumulate, making a diagnosis more apparent, which in hindsight can make an initially missed diagnosis seem to be a diagnostic error. A further limitation is in the calculation of diagnostic error rates, which requires counting the number of diagnostic encounters in the denominator as well as the number of misses in the numerator. Defining a representative sample of diagnostic encounters has been notoriously difficult, making calculating diagnostic miss rates difficult.[3] For example, among patients in an emergency department, defining a representative sample of patients with a possible AMI to calculate a diagnostic miss rate for AMI is difficult. Notwithstanding the difficulties in measurement, diagnostic error remains a substantial problem, and tackling this problem through greater awareness is urgently needed.[13]

Diagnostic reasoning versus management reasoning

Clinical reasoning integrates information on patients and clinical information, medical knowledge, and situational factors to provide care for patients. Clinical reasoning is an umbrella term that includes both diagnostic reasoning and management reasoning. Diagnostic reasoning is a classification task with various levels of specificity (for example, AMI versus STEMI versus STEMI with complete occlusion of the first obtuse marginal artery). Diagnostic reasoning has an objective endpoint, although the gold standard for a diagnosis can have problems of reliability (for example, cardiologists' determination of the presence of congestive heart failure in clinical trials). Management reasoning involves prioritization of tasks, shared decision making with a patient and family, and dynamic monitoring of response to treatment. Management reasoning is more subjective, reflecting context, available resources, and the patient's choice.[17 18] For the purposes of this review, we focus on diagnostic reasoning, acknowledging that this is only a portion of the task of clinical reasoning.

The two step process of making a diagnosis

Cardiac disease is diverse, with a broad range of presenting signs and symptoms, creating diagnostic challenges for the clinician. The diagnosis is often obscured at the time of initial presentation by vague, poorly characterized symptoms such as chest pain, shortness of breath, or fluttering in the chest. The clinician must elucidate the patient's symptoms and translate the patient's own words into the lexicon of cardiology. The clinician then connects signs and symptoms in a recognizable narrative or pattern and works inductively to place the patient's illness in the correct diagnostic category. The process depends on one's ability to engage the patient and elicit a clear and complete history. The importance of the history for making a correct cardiac diagnosis is well summarized by a well known quote attributed to Sir William Osler, "Listen to the patient; he will tell you the diagnosis."

The physical examination can be helpful if it reveals an obvious sign such as a murmur or a friction rub, but often the examination is inconclusive and further diagnostic testing is needed. An organized and selective diagnostic testing strategy is key for proceeding in an effective and efficient manner. Despite skill and experience, the clinician often remains indecisive about a cardiac diagnosis, as evidenced by the fact that about a third of patients labeled with the discharge diagnosis of congestive heart failure were also initially treated for a pulmonary diagnosis.[19]

Clinicians tackle clinical ambiguity and diagnostic uncertainty by using intuition and analytical reasoning. Research from the 1970s showed that the diagnostic process is composed of two parts: hypothesis generation and hypothesis testing, the so-called "hypothetico-deductive method."[20-22]

Hypothesis generation

Researchers have found that expert diagnosticians have between three and five diagnoses in mind within seconds to minutes after starting a diagnostic encounter.[20-23] Generating the hypothesis early in the encounter is important for the accuracy of the eventual diagnosis. In one observational study, if the clinician had the diagnosis in mind early on, the diagnostic accuracy was 95%; if not, the diagnostic accuracy fell to 25%.[24] Another observational study of the diagnostic process showed that primary care physicians can make an accurate diagnosis on the basis of just the chief complaint in 79% of cases.[22] A qualitative study showed that emergency physicians generated 25% of their diagnostic hypotheses before seeing the patient, and they generated 75% of their hypotheses within five minutes of starting the diagnostic encounter.[25] The remarkable ability of experts to recognize diagnostic possibilities was shown to be an effortless and instantaneous use of intuition, or non-analytical reasoning.[26-28] The cognitive psychologist Herbert Simon described how experts use 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.”[29]

Early work showed that the two step diagnostic process was not necessarily restricted to expert diagnosticians but was observed in novice medical students as well.[20 21] The distinguishing feature of the master diagnostician was not possession of a generalizable diagnostic skill, but rather it was the expert’s ability to mobilize and use knowledge from past experience. Moreover, Elstein found that expert diagnostic ability was “content specific.” For both novices and experts, good performance on one case did not translate to good performance on another case of different content. Thus, the accuracy of the diagnostic process was dependent on the clinician’s experiential knowledge. Cardiologists and other specialists seem to know this wholeheartedly. They work within their content area and are quick to seek consultation with others when the diagnosis seems to fall outside of their content expertise.

The expert’s ability to mobilize and use the appropriate knowledge for an accurate diagnosis has led researchers to contemplate the possible knowledge structures that the expert might use to store experiential knowledge.[30] Several possible structures have been proposed, including propositional networks, prototypes, semantic axes, and exemplars.[26 31-33] Research suggests that clinicians are likely flexible in how they encode, access, and mobilize knowledge for various diagnostic encounters.[30]

Studies have shown that the mechanism of diagnostic hypothesis generation varies, depending on a clinician’s amount of experience and level of training.[34 35] Students lack clinical experience and rely on biomedical knowledge to make causal connections to formulate diagnostic hypotheses.[36 37] Observing a student evaluating a patient with chest pain can show how this early method can be slow and relatively ineffective. As trainees gain experience, formal knowledge of basic pathophysiology is combined with expanding clinical experience as their diagnostic competence matures.

Illness scripts
With clinical experience, a student’s knowledge of disease is expanded to include signs, symptoms, and other clinical features that are observed in actual patients. The learner’s biomedical knowledge and growing clinical experience are reorganized and “encapsulated” into narrative structures that are referred to as illness scripts.[35 38-40] The term “script” implies a series of events that, along with enabling conditions, define a knowledge structure for remembering a diagnosis. The presentation of a typical patient with chest pain provides an excellent example of an illness script. Imagine a patient with a family history of coronary artery disease and a smoking history, who presents with a three week history of progressive chest pressure in the mid chest that occurs with exertion and resolves after a few minutes of rest. The enabling conditions, or risk factors, and the classic sequence of events are combined with a basic understanding of coronary artery disease to become encapsulated into memory as an illness script of unstable angina. With this mental representation, a clinician might envision an actual patient combined with a mental image of plaque rupture and myocardial ischemia. An illness script can sometimes incorporate a prototypical patient with typical features, or in other cases it can incorporate a specific patient with particular features. With time, learners accumulate a repertoire of illness scripts that is idiosyncratic to each learner on the basis of his or her individual experience.[35]

Exemplars
With further exposure to actual cases, clinicians gain experiential knowledge by remembering individual diagnostic instances in episodic memory. When an instance or object is labeled, categorized, and placed in long term memory, it is remembered using a knowledge structure called an exemplar.[40 41] Rather than abstracted knowledge, exemplars are direct memories of specific patients with unique features. Exemplars may remain as distinct memories or may become less distinct and more generalized over time. Each clinician has an idiosyncratic patient experience, and the exemplars that are remembered are the result of a clinician’s unique experiences and are not generalizable to other clinicians.

An exemplar can be retrieved from memory effortlessly and unconsciously. The experiential knowledge base of an experienced clinician is analogous to a large file cabinet filled with many exemplars, filed according to diagnostic category. The range and variety of exemplars gives the clinician a sense of the diversity within a diagnostic category and the distinguishing features between diagnostic categories. Clinicians know that exemplars of anterior and posterior myocardial infarctions are within the same category and that an acute aortic dissection belongs to a different category, on the basis of the distinct features of the exemplars. With experience, they are able to quickly recognize the contrasting features of different diagnostic categories, similar to recognizing the contrasting appearance of a right bundle branch block and a left bundle branch block on an electrocardiogram.[42] In addition, expert diagnosticians develop an intuitive sense of the prevalence of a diagnosis based on the number of encounters that are stored in long term memory as exemplars.[43] Expert clinicians know intuitively that an acute myocardial infarction is more common than an aortic dissection on the basis of an implicit sense of the relative number of exemplars in each category stored in long term memory.

Symptom phenotypes

Because exemplars are mental representations of a range of specific clinical symptom constellations (that is, phenotypes), a recent study examined the range of symptom phenotypes in a registry of young patients presenting with AMI.[44] This registry offered an opportunity to study symptom phenotypes of AMI because it prospectively and systematically recorded detailed information about patients' presenting symptoms. Among 3501 patients with AMI, 488 unique symptom phenotypes were identified, showing the degree of variation that challenges diagnosticians. Significantly more symptom phenotypes occurred in women than in men, which might be a source of ambiguity that could help to explain why the diagnosis of AMI is missed more frequently in women than in men.[3] At a population level, the most common symptoms were chest pain, radiation, shortness of breath, and diaphoresis, and these symptoms generally describe the prototypical AMI patient. Interestingly, the phenotype with the prototypical combination of symptoms represented only 1% of the patients in this multicenter cohort. Cognitive psychology studies have shown that many examples are critically important for learning,[45 46] and this study suggests that learners need extensive experience to acquire adequate exposure to various phenotypes to permit the generation of a rich library of exemplars of AMI.

Abductive reasoning

When the patient's presentation is ambiguous and a diagnostic possibility does not intuitively and immediately come to mind, clinicians revert to more reflective reasoning methods.[47] For this, experienced clinicians make use of a thought process called abductive reasoning, or reasoning toward the most plausible hypothesis.[48] Abductive reasoning takes the following course: "The surprising fact C is observed. But if A were true, C would be a matter of course. Hence, there is reason to suspect that A is true."[49] For example, "A patient with a positive troponin is observed. If an AMI is present, troponin would be elevated as a matter of course. Hence, there is reason to suspect an AMI in this patient." Abductive reasoning is a way to work backward and generate hypotheses that might explain observations. It is a form of reasoning that yields hypotheses, not conclusions. It is also the thinking that is used to generate a differential diagnosis by asking, "What other diagnoses could cause the observed findings?" Abductive reasoning describes a method for hypothesis generation that experienced clinicians use when reversion to more reflective thinking is needed.

Hypothesis testing

After rapidly generating several diagnostic possibilities, the clinician begins testing various possibilities, starting with the most likely ones. The decision about which diagnosis to test depends on probability, as well as on the severity and acuity of a potential diagnosis. For example, a lower probability threshold and more urgent testing strategy might be applied for severe life threatening diagnoses such as AMI, pulmonary embolus, or aortic dissection. The prudent strategy might prioritize diagnostic testing for life threatening diseases, but for the most part the order of testing and diagnostic reasoning becomes an exercise in probability.

Bayesian reasoning

The field of cardiology has led the way in thinking probabilistically about possible diagnoses. Clinicians have been shown to use bayesian reasoning to determine the conditional probability of coronary artery disease.[50] Bayesian reasoning provides a method for updating a baseline probability estimate on the basis of the strength of new information. Of course, clinicians rarely, if ever, do formal calculations, but this subjective notion of probability helps the clinician to think about an individual and to use probability estimates to zero in on the correct diagnosis.

No test is perfect, and the strength of new evidence from cardiac testing depends on the degree of imperfection of the test, as measured by the test's operating characteristics. Sensitivity, or the true positive rate, measures the number of patients with disease who test positive; specificity, or the true negative rate, measures the number of patients without disease who test negative. Likelihood ratios can be calculated by combining sensitivity and specificity into dimensionless numbers that give an intuitive estimate of the strength of a positive or negative test result.[9 51]

This approach has limitations too. Measuring sensitivity and specificity requires systematical testing of patients in a research setting, and the numbers, as for any clinical measurements, are relatively imprecise estimates. Furthermore, when the operating characteristics of a test are determined in a rigorously controlled setting and then used in a practice setting or for screening of individuals with a different spectrum of disease, the test's capability can be adversely affected by spectrum bias.[52] For example, a study of the current generation troponin T assay in a rigorous research setting showed that the test had a sensitivity of 95% and specificity of 80%.[53] This research study, however, specifically excluded patients with renal failure or septic shock. When used clinically in an emergency department setting without rigorous restrictions, the false positive rate would likely increase, which would markedly decrease the specificity.[54] Thus, tests should be ordered deliberately to avoid spectrum bias, so as to maximize the operating characteristics of the test and minimize false positive results.

Over time, the prevalence of coronary artery disease, like other cardiac diagnoses, and associated risk factors and local environmental factors, has changed, requiring an updating of the probability estimates.[55 56] Nevertheless, approximating pre-test probability remains an important step in the process of calibrating one's estimate of diagnostic probability.

Anchoring and adjusting

Cognitive psychologists have long recognized that people are not necessarily rational or mathematically rigorous in their decision making.[57-59] People use learned rapid mental shortcuts called heuristics to enable rapid decisions to be made under conditions of uncertainty. Some authors have asserted that heuristics represent a speed-accuracy trade-off, in which the speed of heuristics leads to bias and error.[60] Others have argued that heuristics, although sometimes associated with error, are very useful for rapid decision making.[59 61] One heuristic is anchoring and adjusting,[60] which describes how a decision maker quickly and subjectively estimates (anchors) the baseline probability of an event and then adjusts the probability estimate on the basis of new information. Anchoring and adjusting is an informal shortcut that replaces the formal use of Bayes' rule. It is an intuitive two step process for estimating probability that works in parallel with the two step process of hypothesis generation and hypothesis testing. This heuristic can be affected by two potential biases. One bias is anchoring, whereby the decision maker becomes too stuck on the initial base rate and does not adequately adjust after new evidence is received. The other bias is base rate neglect, whereby the decision maker jumps to a subsequent probability estimate on the basis of new evidence without adequate regard for the initial base rate or disease prevalence.

Our recent experimental study evaluated the effectiveness of teaching the concept of bayesian reasoning to improve the accuracy of diagnostic probability estimates.[62] Students were randomized to receive an instructional video on anchoring and adjusting, likelihood ratios, and bayesian reasoning, versus exposure to repeated examples of cases with feedback, versus no intervention. Previous studies suggested that trainees' subjective probability estimates are often highly inaccurate.[63 64] This study, however, showed that all study participants gave probability estimates that were unexpectedly better than predicted by previous studies. The students who received the conceptual instruction on bayesian concepts showed a modest advantage in estimating the post-test probability of disease, suggesting that even brief instruction on bayesian concepts might improve students' use of statistical heuristics to estimate probability.

Figure 1 shows how a clinician would use the anchoring and adjusting heuristic. The clinician estimates (anchors) a pre-test probability of a diagnosis on the x axis based on knowledge of the base rate or prevalence of a disease. The clinician could draw a vertical line to the curves for either a positive or a negative test result and then a horizontal line to the y axis to determine an estimate of post-test probability. The degree of the adjustment, or the shift in the probability estimate, depends on the strength of a postitive or negative test result, which can be quantified using positive or negative likelihood ratios.

Figure 2 shows how the adjustment or shift in the probability estimate can be asymmetric for different tests. The panel on the left shows a test that is highly specific but not very sensitive, such as a chest radiograph for the diagnosis of congestive heart failure. A positive test result would result in a larger shift in the post-test probability estimate. The panel on the right shows a test that is highly sensitive but not very specific, such as a D-dimer for the diagnosis of pulmonary embolus. A negative test would result in a larger negative shift in the post-test probability estimate.[9]

Dual process theory: system 1 and system 2

The two step diagnostic process is compatible with dual process theory. According to this cognitive psychology theory, two definable systems for thinking exist: one is non-analytical or intuitive thinking, and the other is analytical thinking. Neuroscience

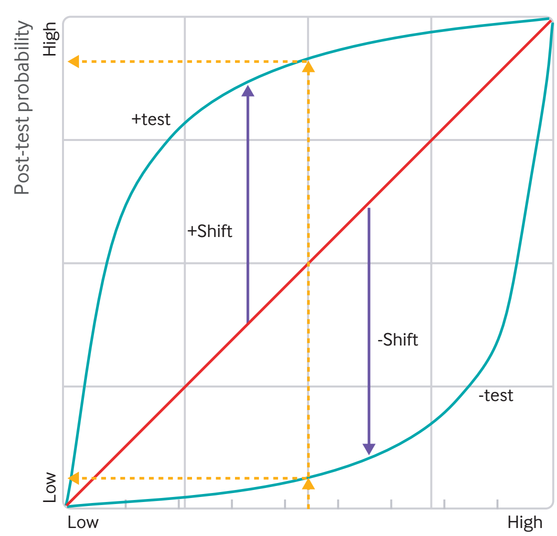

Fig 1 | This graph visually represents how a clinician could use the anchoring and adjusting heuristic. A pre-test probability estimate of about 50% (the anchor) is chosen on the x axis. A vertical line is drawn to the curve for either a positive or a negative test result, and then a horizontal line is drawn to the y axis to determine the post-test probability. The shift shows the degree of the adjustment of the estimate, which depends on the strength of either a positive or a negative test result

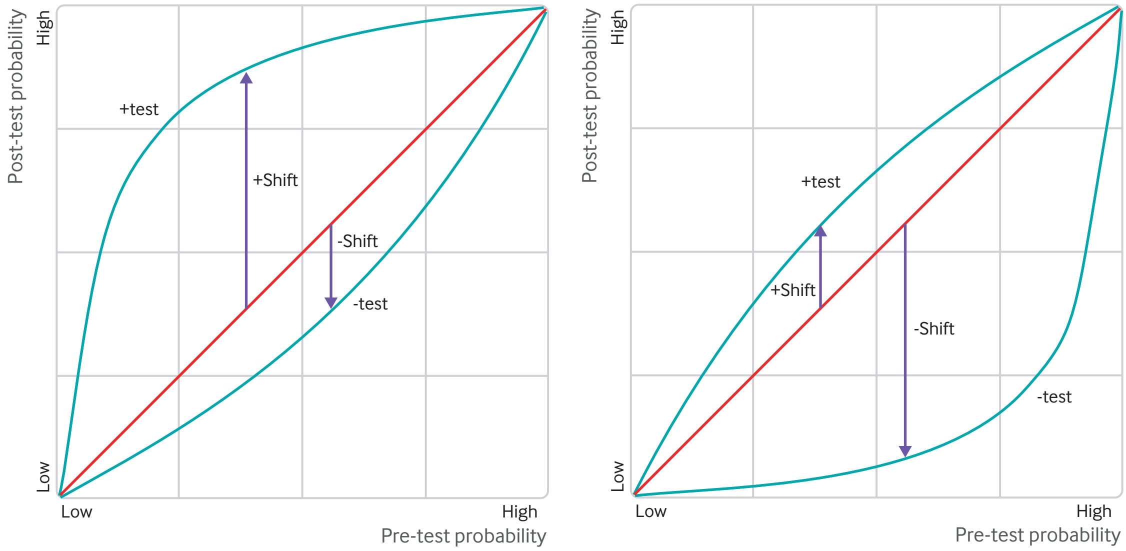

Fig 2 | This graph visually represents how the shift in probability can be asymmetric for different tests. The panel on the left shows a test that is highly specific but not very sensitive, whereby a positive test would result in a larger shift in the post-test probability estimate. The panel on the right shows a test that is highly sensitive but not very specific, whereby a negative test would result in a larger shift in the post-test probability estimate

studies using functional magnetic resonance imaging have shown that these two thinking patterns involve distinctly different areas of the brain and have different metabolic requirements.[65 66] Dual process theory draws a distinction between system 1 thinking, which is intuitive, automatic, quick, and effortless, and system 2 thinking, which is analytic, reflective, slow, and effortful.[58 67 68] System 1 thinking is analogous to driving a car down a familiar, empty highway in that it works unconsciously and effortlessly. System 2 thinking is more like parking a car in a tight parking space, which requires deliberate and effortful attention.

System 1 thinking is triggered by an association between new information and a similar example, or exemplar, stored in long term memory.[69] The association is effortless and depends on the strength of the association, which can be influenced by factors such as the number of examples in memory, the number of common features, and the recency or vividness of the memory.[70] System 2, on the other hand, uses computation, analysis, and logical rules, and places a heavy burden on working memory.[69] Expert clinicians make use of system 1 and system 2 thinking interchangeably, depending on the diagnostic task at hand.

Some authors have promoted the idea that errors occur because of short cuts, or heuristics, which are used by system 1 thinking and not corrected by system 2 reasoning.[58] They provide 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."[58] Other authors have countered this advice by stating that, "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 the same dual process theories have built it into their core terminology."[67]

Does the speed of system 1 thinking lead to diagnostic error? An experimental study showed that a correct diagnosis was actually associated with less time spent on the diagnostic task.[71] In other experimental studies in which investigators cautioned participants about speed and errors and encouraged participants to be deliberate and thorough, these instructions had no effect on diagnostic accuracy.[72-74] In another study, participants were allowed to re-think and revise their initial diagnosis, but revisions were more likely to be incorrect.[75] One study, however, showed that extreme time pressure can have a negative effect on diagnostic accuracy, possibly by inducing anxiety in participants.[76] Most of the evidence suggests that relying less on system 1 and more on system 2, as Kahneman and others advise, does not increase diagnostic accuracy.[77] Kahneman did not study experts, and the context specific knowledge of expert diagnosticians seems to make system 1 thinking more of a strength than a weakness.

Cognitive biases

Some authors have associated diagnostic error with many cognitive biases.[78-83] Surprisingly few studies, however, have empirically examined the role of cognitive biases in diagnostic error. One systematic review of cognitive bias in healthcare indicated that for diagnostic reasoning, only seven biases have actually been empirically evaluated.[84]

The effect of bias on diagnosis can be studied either in an artificial experimental setting or in a practice setting where diagnostic errors have been identified and reviewed retrospectively. One experimental study showed evidence of "satisfaction of search" bias among radiologists shown radiographs with multiple artificial lung nodules. Participants often stopped searching after identifying only one nodule.[85] Another experimental study examined base rate neglect and found little evidence that this bias affects experts.[43] Additionally, these investigators found that the degree of self-reported experience with a diagnosis correlated well with an expert's intuitive estimate of its base rate. Other experimental studies examined availability bias, which refers to how recent exposure to a case can bias the assessment of a new case, and these studies showed mixed results.[86-88] Some experimental studies showed that availability actually enhanced diagnostic accuracy.[89 90] For practitioners, this makes sense. The active engagement of practice with repeated exposures to a range of diagnostic encounters tends to build confidence among practitioners, supporting the notion that availability may improve diagnostic accuracy in the setting of real world practice.

Other observational studies have examined the role of bias through retrospective reviews of diagnostic errors in actual practice. One study examined 100 cases of diagnostic error in the emergency department and found that 68% were associated with a cognitive bias, primarily premature closure.[15] Another observational study, however, found no role of bias in reported cases of diagnostic error.[1] One prospective study showed that diagnostic experts were unable to agree on which bias actually contributed to the diagnostic error, and the study also found that the study participants were themselves affected by hindsight bias.[91]

Strategies for experienced clinicians to avoid diagnostic error

Rather than attributing diagnostic error to biases or flawed cognitive processes, some authors have argued that diagnostic error is more commonly due to an inability to adequately mobilize necessary knowledge, or due to knowledge deficits.[1 77] Experimental studies have shown that the ability to mobilize knowledge and avoid diagnostic errors improves with experience.[86 92 93] Experimental studies have also shown that improved diagnostic accuracy is associated with the acquisition of both formal knowledge and experiential knowledge.[43 71 92]

Simply imploring clinicians to routinely slow down and carefully monitor their intuitive thinking does not seem to be effective for improving diagnostic accuracy. Rapid and intuitive recognition of patterns is an important part of the diagnostic process, particularly in cardiology, and constraining this activity does not seem to be a good strategy. The diagnostic process, however, does allow the opportunity to reflect on the particular features of the diagnostic encounter. Clinicians often ask, "What am I missing? What else could this be?" For tough cases, consciously acknowledging the difficulty of a diagnostic challenge increased accuracy.[94 95] Experienced surgeons seem to know when to slow down at critical moments.[96]

Mobilizing knowledge through deliberate reflection is a promising technique for improving diagnostic accuracy.[97-101] Reflecting on the concordant and discordant features between the patient and the various diagnostic hypotheses offers an opportunity for mid-course correction. Deliberate reflection enables clinicians to overcome distracting and misleading features of a case but requires that the clinician have adequate experience and sufficient clinical knowledge about the diversity of diagnostic features. This has been demonstrated elegantly in a more recent experimental study by these authors.[102] They showed that physicians can be immunized against availability bias through an intervention that increased their knowledge about the features that can discriminate between similar looking diseases. Of course, the strategy of deliberate reflection also requires recognition that an initial diagnostic impression may be unsettled and needs further reflection. Experimental studies have shown that clinicians' diagnostic confidence correlates fairly well with diagnostic accuracy and that the correlation improves with experience, but overconfidence is likely a persistent problem and a potential impediment to the optimal use of deliberate reflection.[94 103]

Checklists have been promoted as tools for improving diagnostic accuracy.[104] Checklists can be classified as either content specific tools that trigger the retrieval of relevant disease specific knowledge or process focused tools that guide adherence to optimal thinking.[105] Unfortunately, studies of the effectiveness of checklists have been disappointing.[105] Checklists focusing on content tend to show some promise in a few limited experimental studies,[106 107] but they tend to be more effective for junior clinicians and for more difficult cases.[105] Whether experimental studies of checklists are generalizable to the practice setting and whether checklists would be used consistently in the real world setting remain open questions.

Computerized decision support programs have also been promoted for improving diagnostic accuracy but have fallen short of expectations.[108 109] A review of a limited number of studies suggested some potential benefit for junior clinicians, but uptake of this technology in practice has been very limited.[110] In part, this may be a consequence of logistical difficulties with the software platform. A recent study of an electronic differential diagnostic support tool showed that computerized decision support can increase the number of diagnostic hypotheses and the probability that the correct diagnosis would be considered, and the impact was greater for novice clinicians.[136]

Cognitive forcing strategies for reducing diagnostic error have been studied.[78 81] Three experimental studies assessed this approach to teach participants to recognize specific cognitive biases and apply such strategies.[111-113] These studies showed that cognitive forcing strategies had no effect on diagnostic errors or accuracy. Humans are not capable of consciously recognizing unconscious biases,[114 115] and that teaching this process would not be successful in reducing errors seems predictable. The task is made more difficult by the fact that more than 100 cognitive biases have been described in the general literature and at least 38 have been described in the medical literature.[116] Moreover, as noted earlier, even experts have trouble consistently and correctly identifying specific cognitive biases.[91]

Notwithstanding difficulties with human introspection, flawed thinking undoubtedly affects diagnostic reasoning. Psychologists have described three general types of fallacies: hasty judgments, biased judgments, and distorted probability estimates.[58] General knowledge of these broad categories might help practitioners to develop habits that help them to avoid bias.

A particular concern is implicit bias regarding race, gender, sexuality, and ability, among other factors. The word bias in this context refers to stereotyping, which is "the process by which people use social categories (for example, race, sex) in acquiring, processing, and recalling information about others."[117] A classic experimental study more than two decades ago showed how race and sex can affect the diagnostic evaluation of chest pain and referral patterns for diagnostic cardiac catheterization.[118] Clearly, implicit bias affects diagnostic reasoning, and overcoming this concern should be central to medical education and professional development. Suggested educational strategies for overcoming implicit bias include emphasizing fairness and egalitarian goals, encouraging identification with common identities with patients, counter-stereotyping, and trying to understand the patient's perspective.[119] More research is needed to identify the most successful educational and support strategies for reducing the adverse effects of implicit bias on the quality of diagnostic reasoning.

Educational strategies to improve diagnostic accuracy

Teaching diagnostic competence in cardiology starts with instruction in the basics of history taking, physical diagnosis, interpretation of electrocardiograms, and a variety of other basic diagnostic skills that are prerequisites for making a cardiovascular diagnosis.[120-122] Professional societies have formulated expanded lists of competencies that are needed for interpretation of cardiovascular diagnostic tests,[123] and these competencies are required for board certification and practice.[124]

True expertise in cardiovascular diagnosis, however, resides in an ability that is learnt through experience and years of deliberate practice and reflection.[125 126] The diagnostic expert uses experiential knowledge gained in the context of training in clinical rotations and specialized practice. Experiential knowledge has been described as "a constantly evolving, dynamic resource, and expertise resides in the ability and willingness not only to use and build, but also to purposefully adapt and re-engineer knowledge effectively."[125]

Several educators and investigators have recommended a variety of educational strategies for promoting diagnostic excellence.[13 115 125-129] The original research in the cognitive science of medical diagnosis was started by educators who were searching for the best way to turn novices into experts.[18-20] Clearly, knowledge, both formal and experiential, is a critical determinant for accurate diagnostic reasoning.[130] That better integration of basic science instruction with clinical experience is a successful strategy for improving diagnostic reasoning is therefore not surprising. Vertical integration of the basic sciences with clinical experience can create cognitive conceptual coherence that seems to improve diagnostic reasoning.[131 132] This strategy may facilitate the formation of illness scripts and make knowledge more accessible at the time of a diagnostic encounter. This integrative strategy may also make basic science education more compelling and memorable because it is linked to relevant clinical context. Contextualizing basic science to specific clinical situations does not necessarily transfer from one content area to another, which may explain why diagnostic expertise is content specific. Most clinical teaching occurs at the bedside, and clinical teachers of session level education need to be aware of the importance of integrating basic science knowledge with experience. Other educational strategies have been described, but measuring educational outcomes is difficult, and relatively few studies have formally evaluated the effect of educational strategies on learning outcomes or have compared recipients of educational interventions with control groups.[130-132]

Societies, professional meetings, and journals have been established to promote educational strategies and interventions with the aim of achieving diagnostic excellence.[133-135] The idea is that greater awareness of the diagnostic process and attention to the sources of diagnostic error could help clinicians to make the most of their experience, purposefully seek feedback, and be more intentional about avoiding diagnostic error. Physicians seek causal explanations and mechanistic concepts, but teaching abstract cognitive psychology concepts out of context may not be effective. Interleaving these cognitive psychology concepts into content specific continuing medical education could be effective, but strategies to improve diagnostic reasoning through continuing medical education need further study.

Conclusion

This narrative review synthesizes the accumulated research into how experts make a diagnosis and considers the implications for learners, clinicians, and teachers. Application of the cognitive science of diagnostic reasoning should help learners to make the most of their clinical experiences and improve the effectiveness of clinicians and teachers. Effective educational strategies are those that focus on the acquisition and mobilization of knowledge, both experiential and formal. Other educational strategies and interventions are promising but need formal study and careful evaluation. Future research should continue to focus on diagnostic reasoning in the context of cardiovascular medicine and other subspecialties, with the goal of improving the accuracy and reliability of the diagnostic process and the quality of care for our patients.

RESEARCH QUESTIONS

- How can researchers design and evaluate real world interventions to assess the effect of implicit bias on the diagnostic process?
- How can teaching of diagnostic reasoning be effectively embedded into subspecialty teaching rotations?
- How can artificial intelligence and computerized decision support tools effectively improve the diagnostic process?

PATIENT INVOLVEMENT

One of the authors reached out to a patient who was thrilled that we were working on this narrative review because his life was dramatically affected by the diagnostic process. The patient and his wife read the manuscript and offered useful comments and encouragement.

The patient was referred to one of the authors with an undiagnosed and severe superior vena cava syndrome. The patient and his wife (a nurse) were very frustrated at the time of presentation because he had been extensively evaluated and the cause of his problem remained undiagnosed. A careful history detailed the enabling conditions and the time sequence of his illness, which led to an explanatory diagnostic hypothesis. This hypothesis led to the confirmatory diagnostic testing and effective treatment. This patient is an example of Osler's adage to "Listen to the patient; he will tell you the diagnosis." This patient remains an inspiration and a reminder of how an accurate medical diagnosis can have an enormous impact on a patient's life.

Contributors: All of the authors contributed to all aspects of the preparation of this review. JEB is the guarantor.

Competing interests: We have read and understood the BMJ policy on declaration of interests and declare the following interests: JEB receives royalties from Dementi Milestone Publishing for the book *The Science of the Art of Medicine: A Guide to Medical Reasoning*.

Provenance and peer review: Commissioned; externally peer reviewed.

1 Zwaan L, Thijs A, Wagner C, van der Wal G, Timmermans DR. Relating faults in diagnostic reasoning with diagnostic errors and patient harm. *Acad Med* 2012;87:149-56. doi:10.1097/ACM.0b013e31823f71e6
2 Moy E, Barrett M, Coffey R, Hines AL, Newman-Toker DE. Missed diagnoses of acute myocardial infarction in the emergency department: variation by patient and facility characteristics. *Diagnosis (Berl)* 2015;2:29-40. doi:10.1515/dx-2014-0053
3 Pope JH, Aufderheide TP, Ruthazer R, et al. Missed diagnoses of acute cardiac ischemia in the emergency department. *N Engl J Med* 2000;342:1163-70. doi:10.1056/NEJM200004203421603
4 Schull MJ, Vermeulen MJ, Stukel TA. The risk of missed diagnosis of acute myocardial infarction associated with emergency department volume. *Ann Emerg Med* 2006;48:647-55. doi:10.1016/j.annemergmed.2006.03.025
5 Studdert DM, Mello MM, Gawande AA, et al. Claims, errors, and compensation payments in medical malpractice litigation. *N Engl J Med* 2006;354:2024-33. doi:10.1056/NEJMsa054479
6 Nallamothu BK, Normand SL, Wang Y, et al. Relation between door-to-balloon times and mortality after primary percutaneous coronary intervention over time: a retrospective study. *Lancet* 2015;385:1114-22. doi:10.1016/S0140-6736(14)61932-2
7 Bradley EH, Nallamothu BK, Herrin J, et al. National efforts to improve door-to-balloon time results from the Door-to-Balloon Alliance. *J Am Coll Cardiol* 2009;54:2423-9. doi:10.1016/j.jacc.2009.11.003
8 Krumholz HM, Anderson JL, Bachelder BL, et al, American College of Cardiology/American Heart Association Task Force on Performance Measures, American Academy of Family Physicians, American College of Emergency Physicians, American Association of Cardiovascular and Pulmonary Rehabilitation, Society for Cardiovascular Angiography and Interventions, Society of Hospital Medicine. ACC/AHA 2008 performance measures for adults with ST-elevation and non-ST-elevation myocardial infarction: a report of the American College of Cardiology/American Heart Association Task Force on Performance Measures (Writing Committee to Develop Performance Measures for ST-Elevation and Non-ST-Elevation Myocardial Infarction) Developed in Collaboration With the American Academy of Family Physicians and American College of Emergency Physicians Endorsed by the American Association of Cardiovascular and Pulmonary Rehabilitation, Society for Cardiovascular Angiography and Interventions, and Society of Hospital Medicine. *J Am Coll Cardiol* 2008;52:2046-99. doi:10.1016/j.jacc.2008.10.012
9 Brush JE. *The Science of the Art of Medicine: A Guide to Medical Reasoning*. Dementi Publishing, 2015.
10 Ladapo JA, Blecker S, Douglas PS. Physician decision making and trends in the use of cardiac stress testing in the United States: an analysis of repeated cross-sectional data. *Ann Intern Med* 2014;161:482-90. doi:10.7326/M14-0296
11 Douglas PS, Hoffmann U, Patel MR, et al, PROMISE Investigators. Outcomes of anatomical versus functional testing for coronary artery disease. *N Engl J Med* 2015;372:1291-300. doi:10.1056/NEJMoa1415516
12 Kohn LT, Corrigan J, Donaldson MS, Institute of Medicine (US) Committee on Quality of Health Care in America. *To err is human: building a safer health system*. National Academies Press, 2000.
13 Balogh E, Miller BT, Ball J, Institute of Medicine (US) Committee on Diagnostic Error in Health Care. *Improving diagnosis in health care*. National Academies Press, 2015.
14 Graber ML. The incidence of diagnostic error in medicine. *BMJ Qual Saf* 2013;22(Suppl 2):ii21-7. doi:10.1136/bmjqs-2012-001615
15 Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. *Arch Intern Med* 2005;165:1493-9. doi:10.1001/archinte.165.13.1493
16 Wachter RM. Why diagnostic errors don't get any respect--and what can be done about them. *Health Aff (Millwood)* 2010;29:1605-10. doi:10.1377/hlthaff.2009.0513
17 Cook DA, Sherbino J, Durning SJ. Management Reasoning: Beyond the Diagnosis. *JAMA* 2018;319:2267-8. doi:10.1001/jama.2018.4385
18 Cook DA, Durning SJ, Sherbino J, Gruppen LD. Management Reasoning: Implications for Health Professions Educators and a Research Agenda. *Acad Med* 2019;94:1310-6. doi:10.1097/ACM.0000000000002768
19 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;64:1574-82. doi:10.1111/jgs.14303
20 Elstein AS, Shulman LS, Sprafka SA. *Medical Problem Solving: An Analysis of Clinical Reasoning*. Harvard University Press, 1978. doi:10.4159/harvard.9780674189089
21 Neufeld VR, Norman GR, Feightner JW, Barrows HS. Clinical problem-solving by medical students: a cross-sectional and longitudinal analysis. *Med Educ* 1981;15:315-22. doi:10.1111/j.1365-2923.1981.tb02495.x
22 Barrows HS, Norman GR, Neufeld VR, Feightner JW. The clinical reasoning of randomly selected physicians in general medical practice. *Clin Invest Med* 1982;5:49-55.
23 Kassirer JP, Gorry GA. Clinical problem solving: a behavioral analysis. *Ann Intern Med* 1978;89:245-55. doi:10.7326/0003-4819-89-2-245
24 Gruppen LD, Woolliscroft JO, Wolf FM. The contribution of different components of the clinical encounter in generating and eliminating diagnostic hypotheses. In: Dabney DS, ed. Research in medical education: Proceedings of the annual conference on research in medical education. Association of American Medical Colleges, 1987:242-7.
25 Pelaccia T, Tardif J, Triby E, et al. How and when do expert emergency physicians generate and evaluate diagnostic hypotheses? A qualitative study using head-mounted video cued-recall interviews. *Ann Emerg Med* 2014;64:575-85. doi:10.1016/j.annemergmed.2014.05.003
26 Norman G, Young M, Brooks L. Non-analytical models of clinical reasoning: the role of experience. *Med Educ* 2007;41:1140-5. doi:10.1111/j.1365-2923.2007.02914.x

27 Brush JEJr, Sherbino J, Norman GR. How Expert Clinicians Intuitively Recognize a Medical Diagnosis. *Am J Med* 2017;130:629-34. doi:10.1016/j.amjmed.2017.01.045
28 Sibbald M, Sherbino J, Preyra I, Coffin-Simpson T, Norman G, Monteiro S. Eyeballing: the use of visual appearance to diagnose 'sick'. *Med Educ* 2017;51:1138-45. doi:10.1111/medu.13396
29 Simon HA. Invariants of human behavior. *Annu Rev Psychol* 1990;41:1-19. doi:10.1146/annurev.ps.41.020190.000245
30 Custers EJ, Regehr G, Norman GR. Mental representations of medical diagnostic knowledge: a review. *Acad Med* 1996;71(Suppl):S55-61. doi:10.1097/00001888-199610000-00044
31 Patel VL, Groen GJ. Knowledge based solution strategies in medical reasoning. *Cogn Sci* 1986;10:91-116. doi:10.1207/s15516709cog1001_4
32 Bordage G, Zacks R. The structure of medical knowledge in the memories of medical students and general practitioners: categories and prototypes. *Med Educ* 1984;18:406-16. doi:10.1111/j.1365-2923.1984.tb01295.x
33 Bordage G. Elaborated knowledge: a key to successful diagnostic thinking. *Acad Med* 1994;69:883-5. doi:10.1097/00001888-199411000-00004
34 Schmidt HG, Norman GR, Boshuizen HP. A cognitive perspective on medical expertise: theory and implication. *Acad Med* 1990;65:611-21. doi:10.1097/00001888-199010000-00001
35 Schmidt HG, Rikers RM. How expertise develops in medicine: knowledge encapsulation and illness script formation. *Med Educ* 2007;41:1133-9. doi:10.1111/j.1365-2923.2007.02915.x
36 Patel VL, Evans DA, Kaufman DR. Reasoning strategies and the use of biomedical knowledge by medical students. *Med Educ* 1990;24:129-36. doi:10.1111/j.1365-2923.1990.tb02511.x
37 Woods NN, Brooks LR, Norman GR. It all make sense: biomedical knowledge, causal connections and memory in the novice diagnostician. *Adv Health Sci Educ Theory Pract* 2007;12:405-15. doi:10.1007/s10459-006-9055-x
38 Feltovich PJ, Barrows HS. Issues of generality in medical problem solving. In: Schmidt HG, De Volder ML, eds. *Tutorials in Problem-Based Learning.* Van Gorcum, 1984: 128-42.
39 Custers EJ. Thirty years of illness scripts: Theoretical origins and practical applications. *Med Teach* 2015;37:457-62. doi:10.3109/0142159X.2014.956052
40 Medin DL. Concepts and conceptual structure. *Am Psychol* 1989;44:1469-81. doi:10.1037/0003-066X.44.12.1469
41 Brooks LR, Norman GR, Allen SW. Role of specific similarity in a medical diagnostic task. *J Exp Psychol Gen* 1991;120:278-87. doi:10.1037/0096-3445.120.3.278
42 Hatala RM, Brooks LR, Norman GR. Practice makes perfect: the critical role of mixed practice in the acquisition of ECG interpretation skills. *Adv Health Sci Educ Theory Pract* 2003;8:17-26. doi:10.1023/A:1022687404380
43 Weber EU, Böckenholt U, Hilton DJ, Wallace B. Determinants of diagnostic hypothesis generation: effects of information, base rates, and experience. *J Exp Psychol Learn Mem Cogn* 1993;19:1151-64. doi:10.1037/0278-7393.19.5.1151
44 Brush JEJr, Krumholz HM, Greene EJ, Dreyer RP. Sex Differences in Symptom Phenotypes Among Patients With Acute Myocardial Infarction. *Circ Cardiovasc Qual Outcomes* 2020;13:e005948. doi:10.1161/CIRCOUTCOMES.119.005948
45 Ross BH. This is like that: The use of earlier problems and the separation of similarity effects. *J Exp Psychol* 1987;13:629-39.
46 Eva KW, Neville AJ, Norman GR. Exploring the etiology of content specificity: factors influencing analogic transfer and problem solving. *Acad Med* 1998;73(Suppl):S1-5. doi:10.1097/00001888-199810000-00028
47 Mamede S, Schmidt HG, Rikers RM, Penaforte JC, Coelho-Filho JM. Breaking down automaticity: case ambiguity and the shift to reflective approaches in clinical reasoning. *Med Educ* 2007;41:1185-92. doi:10.1111/j.1365-2923.2007.02921.x
48 Montgomery K. *How doctors think: Clinical judgment and the practice of medicine.* Oxford University Press, 2006.
49 Hartshorne C, Weiss P, Burks A, eds. *Collected Papers of Charles Sanders Peirce.* Vol 5. Harvard University Press, 1931-58:189.
50 Diamond GA, Forrester JS. Analysis of probability as an aid in the clinical diagnosis of coronary-artery disease. *N Engl J Med* 1979;300:1350-8. doi:10.1056/NEJM197906143002402
51 Brush JE, Krumholz HM. Clinical decision making in cardiology. In: Mann DL, Zipes DP, Libby P, Bonow RO, eds. *Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine.* Elsevier/Saunders, 2015: 25-38.
52 Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *N Engl J Med* 1978;299:926-30. doi:10.1056/NEJM197810262991705
53 Reichlin T, Hochholzer W, Bassetti S, et al. Early diagnosis of myocardial infarction with sensitive cardiac troponin assays. *N Engl J Med* 2009;361:858-67. doi:10.1056/NEJMoa0900428
54 Brush JEJr, Kaul S, Krumholz HM. Troponin Testing for Clinicians. *J Am Coll Cardiol* 2016;68:2365-75. doi:10.1016/j.jacc.2016.08.066
55 Gibbons RJ, Miller TD. Declining Accuracy of the Traditional Diamond-Forrester Estimates of Pretest Probability of Coronary Artery Disease: Time for New Methods. *JAMA Intern Med* 2021;181:579-80. doi:10.1001/jamainternmed.2021.0171
56 Genders TS, Steyerberg EW, Alkadhi H, et al, CAD Consortium. A clinical prediction rule for the diagnosis of coronary artery disease: validation, updating, and extension. *Eur Heart J* 2011;32:1316-30. doi:10.1093/eurheartj/ehr014
57 Simon HA. Invariants of human behavior. *Annu Rev Psychol* 1990;41:1-19. doi:10.1146/annurev.ps.41.020190.000245
58 Kahneman D. *Thinking, fast and slow.* Farrar, Straus and Giroux, 2011.
59 Gigerenzer G, Gaissmaier W. Heuristic decision making. *Annu Rev Psychol* 2011;62:451-82. doi:10.1146/annurev-psych-120709-145346
60 Gilovich T, Griffin DW, Kahneman D. *Heuristics and biases: the psychology of intuitive judgment.* Cambridge University Press, 2002. doi:10.1017/CBO9780511808098.
61 Lopes LL. The rhetoric of irrationality. *Theory Psychol* 1991;1:65-82. doi:10.1177/0959354391011005
62 Brush JEJr, Lee M, Sherbino J, Taylor-Fishwick JC, Norman G. 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. doi:10.1001/jamanetworkopen.2019.18023
63 Casscells W, Schoenberger A, Graboys TB. Interpretation by physicians of clinical laboratory results. *N Engl J Med* 1978;299:999-1001. doi:10.1056/NEJM197811022991808
64 Eddy DM. Probabilistic reasoning in clinical medicine: problems and opportunities. In: Kahneman D, Slovic P, Tversky A, eds. *Judgment Under Uncertainty: Heuristics and Biases.* Cambridge University Press, 1982:249-67. doi:10.1017/CBO9780511809477.019
65 Goel V, Dolan RJ. Explaining modulation of reasoning by belief. *Cognition* 2003;87:B11-22. doi:10.1016/S0010-0277(02)00185-3
66 Masicampo EJ, Baumeister RF. Toward a physiology of dual-process reasoning and judgment: lemonade, willpower, and expensive rule-based analysis. *Psychol Sci* 2008;19:255-60. doi:10.1111/j.1467-9280.2008.02077.x
67 Evans JS. Dual-processing accounts of reasoning, judgment, and social cognition. *Annu Rev Psychol* 2008;59:255-78. doi:10.1146/annurev.psych.59.103006.093629
68 Evans JS, Stanovich KE. Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspect Psychol Sci* 2013;8:223-41. doi:10.1177/1745691612460685
69 Logan GD. Toward an instance theory of automatization. *Psychol Rev* 1988;95:492-527. doi:10.1037/0033-295X.95.4.492
70 Shiffrin RM, Schneider W. Automatic and controlled processing revisited. *Psychol Rev* 1984;91:269-76. doi:10.1037/0033-295X.91.2.269
71 Sherbino J, Dore KL, Wood TJ, et al. The relationship between response time and diagnostic accuracy. *Acad Med* 2012;87:785-91. doi:10.1097/ACM.0b013e318253acbd
72 Ilgen JS, Bowen JL, Yarris LM, Fu R, Lowe RA, Eva K. Adjusting our lens: can developmental differences in diagnostic reasoning be harnessed to improve health professional and trainee assessment?*Acad Emerg Med* 2011;18(Suppl 2):S79-86. doi:10.1111/j.1553-2712.2011.01182.x
73 Ilgen JS, Bowen JL, McIntyre LA, et al. Comparing diagnostic performance and the utility of clinical vignette-based assessment under testing conditions designed to encourage either automatic or analytic thought. *Acad Med* 2013;88:1545-51. doi:10.1097/ACM.0b013e3182a31c1e
74 Norman G, Sherbino J, Dore K, et al. The etiology of diagnostic errors: a controlled trial of system 1 versus system 2 reasoning. *Acad Med* 2014;89:277-84. doi:10.1097/ACM.0000000000000105
75 Monteiro SD, Sherbino J, Patel A, Mazzetti I, Norman GR, Howey E. Reflecting on Diagnostic Errors: Taking a Second Look is Not Enough. *J Gen Intern Med* 2015;30:1270-4. doi:10.1007/s11606-015-3369-4
76 ALQahtani DA, Rotgans JI, Mamede S, et al. Does Time Pressure Have a Negative Effect on Diagnostic Accuracy?*Acad Med* 2016;91:710-6. doi:10.1097/ACM.0000000000001098
77 Norman GR, Monteiro SD, Sherbino J, Ilgen JS, Schmidt HG, Mamede S. The Causes of Errors in Clinical Reasoning: Cognitive Biases, Knowledge Deficits, and Dual Process Thinking. *Acad Med* 2017;92:23-30. doi:10.1097/ACM.0000000000001421
78 Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. *Acad Med* 2003;78:775-80. doi:10.1097/00001888-200308000-00003
79 Croskerry P. Clinical cognition and diagnostic error: applications of a dual process model of reasoning. *Adv Health Sci Educ Theory Pract* 2009;14(Suppl 1):27-35. doi:10.1007/s10459-009-9182-2
80 Croskerry P. A universal model of diagnostic reasoning. *Acad Med* 2009;84:1022-8. doi:10.1097/ACM.0b013e3181ace703

81 Croskerry P. From mindless to mindful practice--cognitive bias and clinical decision making. *N Engl J Med* 2013;368:2445-8. doi:10.1056/NEJMp1303712
82 Redelmeier DA. Cognitive psychology and medical judgment: some downfalls of studying pitfalls. *Med Decis Making* 1991;11:169-70. doi:10.1177/0272989X9101100304
83 Redelmeier DA, Ferris LE, Tu JV, Hux JE, Schull MJ. Problems for clinical judgement: introducing cognitive psychology as one more basic science. *CMAJ* 2001;164:358-60.
84 Blumenthal-Barby JS, Krieger H. Cognitive biases and heuristics in medical decision making: a critical review using a systematic search strategy. *Med Decis Making* 2015;35:539-57. doi:10.1177/0272989X14547740
85 Berbaum KS, Schartz KM, Caldwell RT, et al. Satisfaction of search from detection of pulmonary nodules in computed tomography of the chest. *Acad Radiol* 2013;20:194-201. doi:10.1016/j.acra.2012.08.017
86 Hatala R, Norman GR, Brooks LR. Impact of a clinical scenario on accuracy of electrocardiogram interpretation. *J Gen Intern Med* 1999;14:126-9. doi:10.1046/j.1525-1497.1999.00298.x
87 Schmidt HG, Mamede S, van den Berge K, van Gog T, van Saase JL, Rikers RM. Exposure to media information about a disease can cause doctors to misdiagnose similar-looking clinical cases. *Acad Med* 2014;89:285-91. doi:10.1097/ACM.0000000000000107
88 Mamede S, van Gog T, van den Berge K, et al. Effect of availability bias and reflective reasoning on diagnostic accuracy among internal medicine residents. *JAMA* 2010;304:1198-203. doi:10.1001/jama.2010.1276
89 Allen SW, Brooks LR, Norman GR, Rosenthal D. Effect of prior examples on rule-based diagnostic performance. *Res Med Educ* 1988;27:9-14.
90 Brooks LR, Norman GR, Allen SW. Role of specific similarity in a medical diagnostic task. *J Exp Psychol Gen* 1991;120:278-87. doi:10.1037/0096-3445.120.3.278
91 Zwaan L, Monteiro S, Sherbino J, Ilgen J, Howey B, Norman G. Is bias in the eye of the beholder? A vignette study to assess recognition of cognitive biases in clinical case workups. *BMJ Qual Saf* 2017;26:104-10. doi:10.1136/bmjqs-2015-005014
92 Groves M, O'Rourke P, Alexander H. Clinical reasoning: the relative contribution of identification, interpretation and hypothesis errors to misdiagnosis. *Med Teach* 2003;25:621-5. doi:10.1080/01421590310001605688
93 Norman GR, Rosenthal D, Brooks LR, Allen SW, Muzzin LJ. The development of expertise in dermatology. *Arch Dermatol* 1989;125:1063-8. doi:10.1001/archderm.1989.01670200039005
94 Mamede S, Schmidt HG, Rikers RM, Penaforte JC, Coelho-Filho JM. Influence of perceived difficulty of cases on physicians' diagnostic reasoning. *Acad Med* 2008;83:1210-6. doi:10.1097/ACM.0b013e31818c71d7
95 Mamede S, Hautz WE, Berendonk C, et al. Think Twice: Effects on Diagnostic Accuracy of Returning to the Case to Reflect Upon the Initial Diagnosis. *Acad Med* 2020;95:1223-9. doi:10.1097/ACM.0000000000003153
96 Moulton CA, Regehr G, Lingard L, Merritt C, MacRae H. Slowing down to stay out of trouble in the operating room: remaining attentive in automaticity. *Acad Med* 2010;85:1571-7. doi:10.1097/ACM.0b013e3181f073dd
97 Mamede S, Schmidt HG, Penaforte JC. Effects of reflective practice on the accuracy of medical diagnoses. *Med Educ* 2008;42:468-75. doi:10.1111/j.1365-2923.2008.03030.x
98 Mamede S, Schmidt HG. Reflection in diagnostic reasoning: what really matters?*Acad Med* 2014;89:959-60. doi:10.1097/ACM.0000000000000306
99 Mamede S, van Gog T, van den Berge K, van Saase JL, Schmidt HG. Why do doctors make mistakes? A study of the role of salient distracting clinical features. *Acad Med* 2014;89:114-20. doi:10.1097/ACM.0000000000000077
100 Mamede S, Schmidt HG. The structure of reflective practice in medicine. *Med Educ* 2004;38:1302-8. doi:10.1111/j.1365-2929.2004.01917.x
101 Mamede S, Splinter TA, van Gog T, Rikers RM, Schmidt HG. Exploring the role of salient distracting clinical features in the emergence of diagnostic errors and the mechanisms through which reflection counteracts mistakes. *BMJ Qual Saf* 2012;21:295-300. doi:10.1136/bmjqs-2011-000518
102 Mamede S, de Carvalho-Filho MA, de Faria RMD, et al. 'Immunising' physicians against availability bias in diagnostic reasoning: a randomised controlled experiment. *BMJ Qual Saf* 2020;29:550-9. doi:10.1136/bmjqs-2019-010079
103 Friedman CP, Gatti GG, Franz TM, et al. Do physicians know when their diagnoses are correct? Implications for decision support and error reduction. *J Gen Intern Med* 2005;20:334-9. doi:10.1111/j.1525-1497.2005.30145.x
104 Gupta A, Graber ML. Web Exclusive. Annals for Hospitalists Inpatient Notes - Just What the Doctor Ordered-Checklists to Improve Diagnosis. *Ann Intern Med* 2019;170:HO2-3. doi:10.7326/M19-0829
105 Zwaan L, Staal J. *Evidence on Use of Clinical Reasoning Checklists for Diagnostic Error Reduction.* Agency for Healthcare Research and Quality, 2020.
106 Ely JW, Graber MA. Checklists to prevent diagnostic errors: a pilot randomized controlled trial. *Diagnosis (Berl)* 2015;2:163-9. doi:10.1515/dx-2015-0008
107 Shimizu T, Matsumoto K, Tokuda Y. Effects of the use of differential diagnosis checklist and general de-biasing checklist on diagnostic performance in comparison to intuitive diagnosis. *Med Teach* 2013;35:e1218-29. doi:10.3109/0142159X.2012.742493
108 Berner ES, Webster GD, Shugerman AA, et al. Performance of four computer-based diagnostic systems. *N Engl J Med* 1994;330:1792-6. doi:10.1056/NEJM199406233302506
109 Friedman CP, Elstein AS, Wolf FM, et al. Enhancement of clinicians' diagnostic reasoning by computer-based consultation: a multisite study of 2 systems. *JAMA* 1999;282:1851-6. doi:10.1001/jama.282.19.1851
110 Graber ML, Kissam S, Payne VL, et al. Cognitive interventions to reduce diagnostic error: a narrative review. *BMJ Qual Saf* 2012;21:535-57. doi:10.1136/bmjqs-2011-000149
111 Sherbino J, Yip S, Dore KL, Siu E, Norman GR. The effectiveness of cognitive forcing strategies to decrease diagnostic error: an exploratory study. *Teach Learn Med* 2011;23:78-84. doi:10.1080/10401334.2011.536897
112 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. doi:10.2310/8000.2013.130860
113 Smith BW, Slack MB. The effect of cognitive debiasing training among family medicine residents. *Diagnosis (Berl)* 2015;2:117-21. doi:10.1515/dx-2015-0007
114 Bargh JA, Chartrand TL. The unbearable automaticity of being. *Am Psychol* 1999;54:462-79. doi:10.1037/0003-066X.54.7.462
115 Eva KW. What every teacher needs to know about clinical reasoning. *Med Educ* 2005;39:98-106. doi:10.1111/j.1365-2929.2004.01972.x
116 Elstein AS. Thinking about diagnostic thinking: a 30-year perspective. *Adv Health Sci Educ Theory Pract* 2009;14(Suppl 1):7-18. doi:10.1007/s10459-009-9184-0
117 Smedley BD, Stith AY, Nelson AR, Institute of Medicine (US) Committee on Understanding and Eliminating Racial and Ethnic Disparities in Health Care. *Unequal treatment: confronting racial and ethnic disparities in health care.* National Academies Press, 2003.
118 Schulman KA, Berlin JA, Harless W, et al. The effect of race and sex on physicians' recommendations for cardiac catheterization. *N Engl J Med* 1999;340:618-26. doi:10.1056/NEJM199902253400806
119 Stone J, Moskowitz GB. Non-conscious bias in medical decision making: what can be done to reduce it?*Med Educ* 2011;45:768-76. doi:10.1111/j.1365-2923.2011.04026.x
120 Mangione S, Nieman LZ. Cardiac auscultatory skills of internal medicine and family practice trainees. A comparison of diagnostic proficiency. *JAMA* 1997;278:717-22. doi:10.1001/jama.1997.03550090041030
121 Norman GR, Coblentz CL, Brooks LR, Babcook CJ. Expertise in visual diagnosis: a review of the literature. *Acad Med* 1992;67(Suppl):S78-83. doi:10.1097/00001888-199210000-00045
122 Issenberg SB, McGaghie WC, Gordon DL, et al. Effectiveness of a cardiology review course for internal medicine residents using simulation technology and deliberate practice. *Teach Learn Med* 2002;14:223-8. doi:10.1207/S15328015TLM1404_4
123 Fuster V, Halperin JL, Williams ES, et al. COCATS 4 Task Force 1: Training in Ambulatory, Consultative, and Longitudinal Cardiovascular Care. *J Am Coll Cardiol* 2015;65:1734-53. doi:10.1016/j.jacc.2015.03.023
124 Brush JEJr, Oetgen WJ. Maintenance of Competence in Cardiovascular Training and Practices: Worth the Effort?*Methodist Debakey Cardiovasc J* 2020;16:199-204. doi:10.14797/mdcj-16-3-199
125 Mylopoulos M, Regehr G. Cognitive metaphors of expertise and knowledge: prospects and limitations for medical education. *Med Educ* 2007;41:1159-65. doi:10.1111/j.1365-2923.2007.02912.x
126 Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Acad Med* 2004;79(Suppl):S70-81. doi:10.1097/00001888-200410001-00022
127 Regehr G, Norman GR. Issues in cognitive psychology: implications for professional education. *Acad Med* 1996;71:988-1001. doi:10.1097/00001888-199609000-00015
128 Bowen JL. Educational strategies to promote clinical diagnostic reasoning. *N Engl J Med* 2006;355:2217-25. doi:10.1056/NEJMra054782

129 Schmidt HG, Mamede S. How to improve the teaching of clinical reasoning: a narrative review and a proposal. *Med Educ* 2015;49:961-73. doi:10.1111/medu.12775
130 Norman G. Research in medical education: three decades of progress. *BMJ* 2002;324:1560-2. doi:10.1136/bmj.324.7353.1560
131 Woods NN, Neville AJ, Levinson AJ, Howey EH, Oczkowski WJ, Norman GR. The value of basic science in clinical diagnosis. *Acad Med* 2006;81(Suppl):S124-7. doi:10.1097/00001888-200610001-00031
132 Kulasegaram KM, Martimianakis MA, Mylopoulos M, Whitehead CR, Woods NN. Cognition before curriculum: rethinking the integration of basic science and clinical learning. *Acad Med* 2013;88:1578-85. doi:10.1097/ACM.0b013e3182a45def
133 Graber ML, Wachter RM, Cassel CK. Bringing diagnosis into the quality and safety equations. *JAMA* 2012;308:1211-2. doi:10.1001/2012.jama.11913
134 Singh H, Graber ML. Improving Diagnosis in Health Care--The Next Imperative for Patient Safety. *N Engl J Med* 2015;373:2493-5. doi:10.1056/NEJMp1512241
135 Graber ML, Rencic J, Rusz D, et al. Improving diagnosis by improving education: a policy brief on education in healthcare professions. *Diagnosis (Berl)* 2018;5:107-18. doi:10.1515/dx-2018-0033
136 Sibbald M, Monteiro S, Sherbino J, LoGiudice A, Friedman C, Norman G. Should electronic differential diagnosis support be used early or late in the diagnostic process? A multicentre experimental study of Isabel. *BMJ Qual Saf* 2021;bmjqs-2021-013493.