LEARNING OBJECTIVES

1. The attendee will be able to introduce participants to models of clinical decision-making and potential sources of diagnostic error.

2. The attendee will be able to review diagnostic error in the context of cognitive biases.

3. The attendee will be able to discuss the role of evidence based medicine into clinical decision making in daily practice.

CME QUESTIONS

1. The Availability heuristic refers to:
   a. Paying excessive attention to a piece of information obtained early in the history.
   b. Reaching a diagnosis too quickly and failing to consider other causes.
   c. Making a decision based solely upon potential outcome.
   d. Making a diagnosis based upon easy of recall of previous, similar cases.

2. A patient with known multiple sclerosis presents with ptosis and diplopia due to a right 3rd nerve palsy. She is initially diagnosed with MS relapse, but fails to recover. She is later found to have a right posterior communicating artery aneurysm responsible for the 3rd nerve palsy. Which cognitive biases led to the misdiagnosis?
   a. Outcome bias
   b. Framing bias
   c. Availability bias
   d. Base rate neglect bias

3. The Hawthorne effect refers to:
   a. The tendency for study subjects to behave differently when they are being observed.
   b. A tendency to enjoy early American writers.
   c. The tendency for physicians to stop searching once an apparent diagnosis has been reached.
   d. The tendency to over-rate our diagnostic acumen.

KEYWORDS

1. Cognitive Biases
2. Clinical Reasoning
3. Diagnostic Error
4. Heuristics
5. Decision-Making

Evidence-based medicine (EBM) seeks to improve patient care and clinical outcomes by promoting the integration of research-driven evidence into clinical decision-making (CDM). An often unstated prerequisite to effective practice of EBM is that the physician possesses strong clinical reasoning skills. In fact, there is often a dense fog of clinical uncertainty through which a physician must navigate before reaching a point where EBM can more directly guide decision making. For example, the literature provides relatively little guidance regarding how to perform a history and clinical examination on a patient presenting with diplopia and blurred vision. However, EBM can provide data to indicate whether MRA, CTA, or catheter angiography is best able to detect an expanding posterior communicating artery aneurysm resulting in a third nerve palsy. At its core, EBM rests upon foundational evidence (e.g., biological plausibility), and allows for integration of best available evidence with clinical judgment and expertise.

All physicians rely upon clinical reasoning skills when evaluating patients. The growth in population and available medical technology (increased survival, transplant, HAART therapy, etc) with increased attention on resource utilization means that the average physician often finds himself or herself balancing quick and accurate diagnosis of challenging diagnostic dilemmas against an overlooked vision or life-threatening illness, and therefore ordering (potentially) unnecessary tests. More sophisticated, and necessarily more time consuming and expensive, diagnostic techniques compel us to be even more precise with clinical diagnosis and rational clinical decision-making. Even when imaging is necessary, it is critical to know what image to order (CT or MRI), how to order (with or without contrast, orbital views with fat saturation), and where to image (brain, orbit, cervical spine, skull base). In general, the wider the diagnostic net is cast, the higher the chance that unrelated positive findings may surface to further muddy the waters. Every diagnostic test, including blood work,
imaging, and even histopathology, has a false positive rate, and every positive findings needs to be matched to the patient, particularly if it doesn’t fit the clinical scenario. For example, an MRI of the brain is performed on a patient presenting with unilateral ptosis, and shows a large ipsilateral parasellar meningioma. That is clearly a positive finding- but does it have anything to do with the ptosis? If the ptosis is ultimately due to levator dehiscence, the tumor is an incidental finding, and there may be no need for intervention upon an asymptomatic benign tumor.

Consider a typical case presentation: a 34 year old woman presents with a 4 day history of visual loss in the left eye. She reports pain on eye movements, and previous episodes of leg numbness and gait imbalance. Even before performing the examination, most neuro-ophtalmologists would consider demyelinating optic neuritis as the most likely diagnosis. However, what if the patient were HIV+, with a low CD4 count? Or a 75 year old woman? Or had a known history of previous functional visual loss? Or had a strong maternal family history of bilateral visual loss? These elements would factor into the differential diagnosis- but could also potentially influence the interpretation of physical examination and laboratory findings.

DIAGNOSTIC ERROR AND CLINICAL DECISION MAKING
Diagnostic errors in medicine have become the focus of increased attention over the past decade. They may include medication errors, surgical mistakes, skill deficiencies, and misdiagnoses.

Diagnostic errors comprise a substantial and costly fraction of all medical mishaps. Blendon and colleagues surveyed patients and physicians, and found that 35% of physicians and 42% of patients reported that they or a family member had experienced medical error. Another, larger survey of 2201 adults in the US found that 35% had experienced a medical mistake in the past 5 years; of these, half were described as diagnostic error, and 35% resulted in permanent harm or death. In one study diagnostic errors represented the second largest cause of adverse events. They are the leading or second leading cause of malpractice claims in the US and abroad. One study of autopsy results found diagnostic discrepancies in 20% of cases- of these, in approximately half would the correct diagnosis have changed management.

It should be noted that patients may not always interpret adverse events accurately, or may differ from physicians as to the cause of the unforeseen outcome. Berner and Graber published an extensive review of diagnostic error, and found that error rates differed by clinical specialty. They were lowest in the “perceptual” specialties, radiology and pathology, ranging from 2-5% for both. In contrast, the diagnostic error rate for emergency departments is higher, up to 12% in one study.

There is an existing literature regarding diagnostic error in ophthalmology and neurology, although this is relatively understudied compared with larger specialties such as internal medicine. In most of these studies, observers examine clinical images (fundus photographs, OCT, MRI, etc) and are asked to grade the findings, or compare with a gold standard. There are a few recent studies specifically looking at diagnostic error per se, but fewer still evaluating the specific role clinical reasoning plays in such errors.

Graber published the largest systematic review of diagnostic error to date. He reviewed 100 cases of diagnostic error identified through autopsy results, quality assurance activities, and voluntary reports. 90 cases involved injury, with 33 deaths. He identified three broad categories of error: No-fault errors, in which the nature of the case presentation (including patient compliance) precluded diagnosis; system errors, reflecting latent flaws in the health care system; and cognitive errors, in which inaccurate clinical reasoning or faulty data gathering played a role. Cognitive errors were involved in 74% of cases. These findings suggested that diagnostic error was often multifactorial, involving system-related and cognitive factors.

Most diseases do not present in a classic fashion. Similarly, most patients are not rigorously compliant with history, recollection, and follow up. In the Case Records of the Massachusetts General Hospital, published periodically in New England Journal of Medicine, a case is presented to an expert clinician, who is challenged to identify the correct diagnosis. The case is reasoned through in a classic, analytic method of clinical reasoning. Even so, in all cases presented from 1989 to 1996, the diagnosis was missed by 25% of discussants. In some cases, the disease presented in an atypical fashion, and in others, the disease was newly discovered or a new variant.

THE INFLUENCE OF COGNITIVE BIASES IN CLINICAL DECISION MAKING
Medical diagnosis is an example of decision making under uncertainty. It begins with perception: the clinician identifying physical findings, the interpretation of laboratory tests, review of clinical images (OCT, fundus photographs, MRI, etc.). Although gaps in knowledge play a role in diagnostic error, and are emphasized in the popular press, Graber and others have suggested that cognitive processing errors (or cognitive biases) are far more common than knowledge gaps.

The root cause of such errors can be challenging to determine- error is often identified “after the fact”, and through retrospective analysis. Most physicians do not receive formal instruction in clinical reasoning during their training, and while a physician often recognizes when they have made a mistake, they may not completely understand the cognitive processes underlying their error.
There is no universally accepted model that fully explains the cognitive underpinnings of clinical reasoning. Figure 1 illustrates two similar methods of arriving at a diagnosis and management plan\(^2\). Elstein\(^23\) described the “hypothetico-deductive model” of clinical decision-making. We reduce uncertainty in a clinical case by generating one or more diagnostic hypotheses and procure additional information to confirm or exclude one or more of these hypotheses. The problem is placed in a deductive framework: if patient has X, then he or she must exhibit these features.

This type of reasoning, also called “analytical reasoning” implicitly or explicitly incorporates Bayes’ theorem or regression analysis\(^24,25\). Briefly, these models assume that each clinician is cognizant of the a priori probability with which a particular diagnosis may present, and the conditional probability associating each bit of evidence (clinical findings, presenting symptoms, diagnostic tests) with the diagnosis. Each diagnosis or treatment is assigned a probability that it is the correct choice. These are considered “prior probabilities.” We then collect information that will best test the competing hypotheses. Post-test, or “posterior”, probabilities are then calculated for each diagnosis or treatment under consideration in light of the new finding (including historical and physical exam results) or test result. This process continues serially until a diagnosis or treatment decision is reached.

The advantages of this “forward-flow” approach seem obvious: a problem is encountered (blurred vision, ptosis, diplopia, etc)- a series of plausible hypotheses are generated (optic neuritis, third nerve palsy, etc), and information is gathered in a systematic fashion (duration, associated findings, neuroimaging results, etc) to confirm or refute the original hypothesis.

However, there are potential drawbacks to this approach. Although this model represents an idealized method of moving from clinical uncertainty to a definitive diagnosis, the application is fraught with difficulty. First, there is a heavy reliance upon working memory, which has real limitations in speed and capacity. Second, this model requires large amounts of highly specific information that may not be readily available, such as accurate estimates of prior probabilities. There are also issues with diagnostic tests, which have variable sensitivities and specificities. Other issues include\(^22\): 1) generating a comprehensive differential diagnosis from a paucity of signs and symptoms; 2) predicting which test is likely to have the greatest diagnostic value; 3) recognizing multiple disease processes in a single patient; 4) manipulating multiple probabilities without the assistance of a computer; and 5) moving through this process in a timely fashion.

In daily clinical care settings, most physicians rely upon what some have called “non-analytic” clinical reasoning, or less technically, pattern recognition\(^10\). This involves the use of medical heuristics, or “rules of thumb”, which are derived from time spent in training and years of clinical experience. For example, when confronted with a patient with ptosis, we are generally able to make a diagnosis of levator dehiscence within a short period of time, with minimal diagnostic challenge. This is because we have seen many cases of ptosis, are familiar with the presentation of a variety of conditions which cause ptosis, and are “tuned in” to red flags in the history or exam which would suggest an alternate, more serious, diagnosis. The power of heuristics is enormous, allowing us to navigate the diagnostic challenges in a busy working practice on a daily basis, making effective, usually accurate diagnoses in real time. Croskerry\(^21,26\) refers to this process as “Flesh and Blood Decision-making”, clinicians responding to time pressure and resource availability to make expeditious decisions and institute management plans. The “Casablanca” strategy is an example of Flesh and Blood Decision-making: the usual suspects are rounded up (in terms of blood work, imaging, etc) which allows for additional time for events to mature, decline, or otherwise declare themselves.

One premise of cognitive psychology (a branch of psychology which examines how people reason, formulate judgments, and make decisions) is that errors in diagnostic reasoning are due to predictable categories of cognitive errors (or biases) to which all humans are prone. Biases are defined as inaccurate beliefs that affect decision-making and serve as sources of error\(^20\).

The following table summarizes some of most commonly described biases influencing CDM\(^17,20-22\),
**Step in Diagnosis** | **Bias** | **Definition** | **Example**
--- | --- | --- | ---
**Generating a differential diagnosis** | Availability | Ddx influenced by what is easily recalled (false prevalence) | Diagnosing optic neuritis in a patient presenting with blurred vision and pain around the eye
| Representativeness (judging by similarity) | Suspicion influenced only by s/sx; neglects prevalence of other conditions | | 
| Framing | Being swayed by the wording to focus on certain aspects of a case more than others | | 
**Weighting items in ddx Selecting dx Validating dx** | Confirmation | Additional testing confirms suspected dx but fails to test competing hypothesis | Patient with diplopia has MRI showing parasellar tumor, but true cause is decompensated strabismus
| Anchoring and Adjustment | Failure to adjust ddx based on new data; adherence to original hypothesis | | 
| Search Satisfaction | Stop searching for additional dx once initial dx confirmed | | 
| Blind Obedience | Over-reliance on laboratory testing or imaging | | 
**Choosing course of action (intervention)** | Outcome | Decision judged by outcome rather than logic | Treating patients with corticosteroids (for example) when disease has no known treatment
| Omission | Emphasis on avoiding effects of Rx; under-utilization of treatment | Not treating patient with suspected GCA due to fear of steroid-related complications

**Framing** refers to the error of initiating diagnostic reasoning by overvaluing an item of clinical history presented early in the process\(^{17,27}\). It can be a very useful heuristic to help focus thinking and allow efficient clinical decision making. It can also bias the physician too heavily toward one diagnosis to the exclusion of others. For example, a patient presenting with acute visual loss reports a history of similar visual loss in his mother—this would appropriately raise suspicion of Leber’s Hereditary Optic Neuropathy, but the clinician might be less likely to consider other, far more common, causes of acute visual loss. A large body of evidence, both from medicine and social sciences, suggests that how information is framed greatly influences decision-making\(^{27-29}\). Framing effects can influence both physicians and patients. For example, how we frame management options—even if we believe we are being impartial—can sway patients toward one treatment or the other\(^{28}\).

Graber\(^{20}\) and others\(^{21,25}\) have suggested that “**premature closure**” (which would include Confirmation, Anchoring and Adjustment, and Search Satisfaction described above) is the most common error resulting in faulty CDM. These errors are only rarely the result of simple sloppiness or taking shortcuts— they result from distorted perceptions and under recognized susceptibility to these biases. Graber’s study\(^{18}\) of diagnostic error in internal medicine found that incomplete history taking and physical examination, failure to consider the correct diagnosis, and bias toward a single explanation are all correlated with premature closure. The latter (bias toward a single, unifying explanation) deserves...
special mention. Most of us are trained to endorse medical parsimony, or "Occam’s Razor": the belief that clinicians should favor the simplest hypothesis capable of explaining a set of observations. This heuristic is famously observed and in many cases practical and useful. However, a perhaps less well-known counter-argument is Hickam’s Dictum: “patients can have as many diseases as they please”. This principle is commonly attributed to John Hickam, the former chairman of Medicine at Duke University. This dictum simply asserts that at no stage should a particular diagnosis be excluded solely because it doesn’t appear to fit the principle of Occam’s razor. The principle of Occam’s razor, or parsimony, does not demand that we necessarily opt for the simplest explanation, but instead guides us to seek explanations, without unnecessary additional assumptions, which are capable of accounting for all relevant evidence.

**Representativeness** (a form of pattern recognition) is one of the most commonly used and valuable heuristic for clinicians. If a clinical presentation (a young woman with painful visual loss and ipsilateral RAPD) resembles a well-defined disease (demyelinating optic neuritis), then it is likely that the patient has that disease. However, one weakness of this heuristic is the tendency to ignore prior probabilities and base rates of disease, particularly in complex cases. Some experiments in social psychology illustrate this principle. An example from neuro-ophthalmology would be suspecting a rare disease (say, CNS vasculitis, or paraneoplastic optic neuropathy—or a “Walsh wheel” diagnosis) when an atypical presentation of a more common disease (such as microvascular cranial neuropathy, or NAION) is correct.

Another source of error is **overconfidence**—defined in this sense as miscalibration of one’s own sense of diagnostic accuracy and actual accuracy. This miscalibration can be easily demonstrated in experimental settings, nearly invariably in the direction of overconfidence. Novices in any discipline routinely overrate their skills. This holds for medical trainees as well—in one study, residents performed worse than faculty physicians, but were more confident in the correctness of their diagnosis. In surveys of academic professionals, 94% rate themselves in the top half of their profession. One could only wonder what such a survey of neuro-ophthalmologists would show! In the autopsy study described previously, in which the discordance rate between pre-mortem and post-mortem diagnosis was 20%, physicians were asked to provide the clinical diagnosis and their level of medical certainty. Clinicians who chose the highest level of certainty were wrong 40% of the time. Similar finding were found by Landefeld et al: the level of physician confidence showed no correlation with the ability to predict accuracy of diagnosis. The inability of humans to accurately judge what they know is pervasive, and can be seen in many areas and many types of tasks. When a physician becomes wedded to his or her initial diagnosis, this can result in failure to consider alternate hypotheses.

**Omission and outcomes** biases adversely impact treatment decisions and impede the practice of EBM by focusing the clinician (again, consciously or unconsciously) and the patient on what could happen, instead of what is most likely to happen, following an intervention. Studies of omission bias indicate that physicians underutilize preventive measures to avoid having a direct role in bad outcomes. Outcome bias discounts the evidence underlying medical decisions and places heavy emphasis upon patient outcomes. This is magnified by the very human tendency to want to “do something”, even if doing something may result in the same (or worse) outcome as doing nothing (beyond follow up, observation, and support for the patient). Since many conditions in neuro-ophthalmology have no proven treatment (NAION, Leber’s, microvascular cranial neuropathy) this can result in the use of unproven and potentially harmful therapies, simply because “something must be done”. Outcome bias may result in the under- or over- estimate of a given therapy’s efficacy. For example, if a specific condition can spontaneously improve (true for NAION, optic neuritis, to some degree retinal vascular occlusion), then using an unproven treatment may lead to the erroneous assumption that the treatment “worked”, which can influence the management strategy for future patients with this disease.

**BARRIERS TO GENERATING ACCURATE CAUSAL INFERENCES**

Humans are built to make rapid causal inferences. We have evolved to be skilled, pattern-seeking, causal-finding creatures. At the most basic level, causal inferences are the ability to connect consequences with actions: Y follows X, therefore Y was caused by X. From a pragmatic standpoint, causal inferences are the mechanism we use to decide on a day to day basis what is good (or safe) for us, and what is not. We continuously conduct miniature experiments on ourselves (or participate in experiments driven by social and physical environments), and observe the results. Unfortunately, patterns are difficult to recognize, in the sense that they are hard to differentiate from simple coincidences. Since we are “hard-wired” to see patterns, that sometimes means we see patterns where no pattern exists.

In simple systems, causation is fairly easy to infer. A child who burns his or her hand on a stove generates a rapid, accurate and powerful causal inference which informs future decision-making. However, in complex systems, generating an accurate causal inference is much more difficult. There are several scenarios in which determining an accurate causal inference is challenging:

1. The consequence does not always occur as a result of the action (for example, adverse drug reactions).
2. The consequence does not appear immediately (some treatments have a delayed effect).
3. The consequence can occur in the absence of any action (natural history of the condition includes spontaneous improvement).

4. The consequence is measured along a subtle gradient (for example, visual field progression in IIH or glaucoma).

5. The consequence has a large subjective component (pain, visual acuity).

Any time we treat a patient, we are conducting a small scale experiment. We institute Treatment A and see result B- ipso facto, Treatment A caused Result B. However, this causal inference may not be valid or accurate- there may be other factors which influence the development of Result B, some of which we may not even be aware. More perniciously, even if we ARE aware of these factors, we may fail to account for them due to cognitive biases the cloud our perceptions. Some of these cognitive distortions are well-recognized:

Selective Memory: we tend to recall favorable results more easily than unfavorable.

Response bias: patients are more likely to report positive results than negative ones- i.e., the patient may not report worsening symptoms due to a desire not to displease the physician.

Attrition (“dead men tell no tales”). Patients with unfavorable outcomes may simply not return, and not register in our memory.

The Hawthorne Effect: this refers to a well-described phenomenon in which study subjects behave differently when they know they are involved in a treatment trial as if there weren’t enough to confound results, these subjects have more of a tendency to tell the experimenter what they want to hear. Although this phenomenon is described in relation to large scale treatment trials, there is some evidence that this phenomenon may occur in the smallest scale treatment trial- treating an individual patient with a novel therapy.

All of these factors converge to impede our ability to generate correct, accurate causal inferences, and may influence our clinical decision making.

Potential Solutions:

Much of literature regarding medical decision-making concentrates on primary and emergency care. The degree to which these cognitive biases influence sub-specialist decision making is understudied. However, in some ways, super-sub-specialists (such as neuro-ophthalmologists) may be even more prone to certain errors. For example, we often see patients who have already seen by a primary ophthalmologist or neurologist. There may be an unconscious assumption that the patient has already been “screened” for non-neurologic conditions such as refractive error, cataract, uncompensated congenital strabismus, etc. In many cases, this turns out not to be the case, and it becomes critical to keep common disorders in mind, even in a tertiary or quaternary referral clinic. Subspecialty clinics also tend to see rare diseases more frequently, and one can easily fall prey to a cognitive bias called base rate neglect-thinking every case is a “zebra” (or a unicorn), and spending an enormous amount of time and money to prove that it is, indeed, just a horse.

On the other hand, sub-specialists may have a higher level of “clinical expertise”, a term that is admittedly somewhat nebulous. Research examining expert decision-making (in all fields) has concluded that rapid and accurate pattern recognition is characteristic of experts. Gigerenzer and Goldstein suggest that most real world decisions are made using “automatic” skills, with fast and frugal heuristics that often lead to correct decisions. Expert thinkers in most disciplines are highly valued, as they make rapid decisions, but are ready and willing to reconsider the original hypothesis if the data does not fit. Therefore, they successfully integrate and incorporate analytic and non-analytic methods of decision-making. The challenge is identifying, developing and allocating this “clinical expertise” in a rational manner.

In some cases, simple awareness of our own biases and the manner is which we reason through cases may be sufficient. This process, termed metacognition, implies a reflective approach to clinical problem-solving- taking a step back from the problem at hand and examining our own thought processes. Croskerry and others have suggested strategies such as forced consideration of alternative possibilities, simulation training (cognitive walkthrough strategies), and cognitive aids to decrease reliance on memory (hand held computers, mnemonics, etc). How these strategies would be integrated into a busy clinical practice remains uncertain. Some have advocated for a “cognitive autopsy” after any diagnostic error- a detailed self-review of where clinical decision-making broke down in a particular case.

Redelmeir has suggested some corrective strategies for some specific cognitive biases:
One option that already exists to perform such an “autopsy” is Morbidity and Mortality (M & M) conferences. Originally designed to provide a forum for physicians to review their mistakes and help colleagues from making similar one, most M & M sessions mirror individual response to failure: avoidance, blame, and denial. One review found that only 25% of M & M sessions studied dealt directly with failure. Restructuring M & M conferences reasoning could be another approach to reduce diagnosing between analytic and non-analytic strategies- judiciously using the powerful heuristics we have developed through years of training and clinical practice, yet remaining mindful of some of the potential drawbacks inherent in relying solely upon pattern recognition. Neuro-opthalmologists, as sub-sub-specialists, are in a unique position: we have potentially developed enough clinical expertise to navigate the murky waters of complex cases and streamline diagnostic testing, but may be prone to certain cognitive biases that influence decision-making. Future research is likely to focus on computer-bases decision-making tools and increasing feedback as methods of reducing diagnostic error and improving resource utilization.

### CME ANSWERS

1. D
2. B
3. A

### REFERENCES:

45. Pierluissi E, Fischer MA, Campbell AR et al. Are medical errors discussed in morbidity and mortality conferences? JAMA 2003;290:2838-2842