

Horses and zebras: probabilities, uncertainty, and cognitive bias in clinical diagnosis



Emmet Hirsch, MD

Medical diagnosis is like an iterative logic puzzle in which a list of candidate solutions is winnowed down through step-wise integration of new data, excluding conditions that are not consistent with the facts and retaining those that are. We usually rank the remaining possible diagnoses according to probability. The latter process of odds estimation is summarized creatively in the old aphorism familiar to most physicians trained in North America: “If you hear hoofbeats, think horses, not zebras.”

Considerations other than probability also impact the diagnostic process, such as the risks, benefits, costs, inconvenience of testing, and the potential harmful impact of a missed diagnosis. For example, we might suspect, based on symptoms and physical examination, that a female patient has a simple urinary tract (bladder) infection. We send a culture to confirm the diagnosis, often prescribing antibiotics empirically while awaiting the laboratory result. The patient usually will improve, sometimes because of treatment, sometimes through spontaneous resolution of the infection, and sometimes because she never had a urinary tract infection in the first place. Either way, nothing further is required at the initial visit. The likelihood of a correct diagnosis is high enough and the risks due to error low enough to justify a limited workup.

It is important, however, to hold onto the list of differential diagnoses, in case the patient does not respond as expected or evidence emerges that does not fit our working hypothesis. If the urine culture is negative and the patient fails to improve, we should consider alternatives, such as stone, kidney infection, interstitial cystitis, cancer, trauma, and a fairly long list of other possibilities, depending on the clinical features. There are usually plenty of deer, elk, caribou, and even goats to consider before resorting to “zebra.”

But even an exceedingly rare diagnosis is sometimes the correct one. You can read about such cases nearly every week in both medical and lay sources. Diagnostic conundrums often revolve around correctly identifying the zebra.

And yet, there are occasions when we diagnose “zebra,” ignoring the horse galloping in plain sight.

A colleague once cared for a woman who came to the emergency room with unilateral abdominal pain. The date of the last menstrual period was unknown, and the plasma concentration of the pregnancy-specific hormone human chorionic gonadotropin was 900 mIU/mL. An ultrasound report, ordered by the emergency room staff and resulted before the gynecology team arrived to assess the patient, stated that a cyst within the ovary was “suggestive of ectopic pregnancy,” without mentioning other possibilities. The patient’s discomfort persisted, and she was taken to the operating room, where the cyst, bleeding lightly from its surface, was excised.

No one on the team recognized that rather than an ovarian ectopic, the patient was far more likely to have an early intrauterine pregnancy with a corpus luteum cyst of the ovary. Ovarian ectopic pregnancy, unlike tubal ectopic pregnancy, is exceedingly rare (as few as 1 in 40,000 conceptions¹). This and 2 additional crucial facts were well-known to the attending gynecologist and resident staff: (1) the human chorionic gonadotropin concentration of 900 mIU/mL was below the “discriminatory zone,” the value at which the absence of ultrasound evidence of intrauterine pregnancy can be used to suggest ectopic pregnancy.² With existing technology, the discriminatory zone threshold stands at approximately 1500 mIU/mL. Thus, an intrauterine pregnancy should not have been ruled out in this case. (2) A corpus luteum is present in 100% of normal early pregnancies. In fact, the corpus luteum’s production of progesterone is necessary to sustain the embryo through the first 9 weeks of gestation.³ Failure to recognize the true diagnosis persisted in the operating room, despite the observation, according to the operative report, that the cystic structure contained “yellowish material resembling fat” (ie, the classic appearance of a corpus luteum). The patient’s consequent progesterone deficiency was not replenished, and she miscarried shortly thereafter.

What were the cognitive errors that led to favoring a rare (and erroneous) diagnosis over the likely one? One factor seems to be the phenomenon known in behavioral psychology as “anchoring bias” (unwillingness to modify an initial determination). Once the team had fixed on ovarian ectopic pregnancy, they did not relinquish that diagnosis, despite the persuasive contrary evidence already in hand.

A second potential source of error was “overconfidence bias” (unjustified faith in the reliability of one’s information, skills, or judgment). Specifically, the team may have placed excessive confidence in the imaging report. Pelvic examination is an imprecise tool; in modern gynecology, there is almost no intraabdominal pelvic complaint that does not include an ultrasound scan as part of the workup. Although

From the Department of Obstetrics and Gynecology, NorthShore University Health System, Evanston, IL, and the Department of Obstetrics and Gynecology, University of Chicago Pritzker School of Medicine, Chicago, IL.

The author reports no conflict of interest.

Corresponding author: Emmet Hirsch, MD. ehirsch@northshore.org

0002-9378/\$36.00 • © 2020 Elsevier Inc. All rights reserved.
<https://doi.org/10.1016/j.ajog.2020.01.010>

ultrasound scanning often contributes to optimal management of gynecologic cases, its specificity is limited. Had the radiologist mentioned alternatives to ovarian ectopic pregnancy in the report, the surgeons might have questioned that diagnosis. Indeed, they might have avoided an operative procedure altogether, because the patient's pain was not severe, nor was she hemodynamically unstable.

Recognition of cognitive bias (systematic deviation from purely rational considerations) and how it affects decision-making originated with psychologists Daniel Kahneman and Amos Tversky in a series of groundbreaking papers published primarily in the 1970s. Kahneman and Tversky's work was recently popularized in the book *The Undoing Project* by Lewis⁴ and in Kahneman's own best-selling book, *Thinking Fast and Slow*.⁵ Kahneman won the Nobel Prize in Economics for this research (Tversky died before the granting of the award, which is not given posthumously). Cognitive biases (also known as "heuristics" or mental shortcuts) include anchoring, overconfidence, availability (preferring easy-to-obtain or recently experienced evidence over less readily available or more remote evidence), sunk costs (persistence in pursuing false trails because of effort and expenditures already made), and others. Clinicians, like most people, are subject to these and other sources of error.

What can we do to minimize mistakes that result from cognitive biases? The most important intervention is perhaps the simplest: being mindful of these biases and acting intentionally to counter them, including deliberate consideration of alternatives. The systematic use of tools such as checklists and "hard stops" provides an opportunity to avoid errors of both omission and commission.⁶

The National Academy of Medicine (formerly the Institute of Medicine), in its exhaustive examination of diagnostic processes,⁷ endorsed a team-based approach with liberal use of consultation. They further point to the untapped potential of health information technology tools that can help discern hard-to-identify patterns that are obscured by the myriad data points in a patient's care record. Although it is easy to vilify electronic patient care systems, one of their strengths is immunity to human biases. There is growing evidence, for example, that computer algorithm-based alerts improve accuracy and diminish delays in diagnosis.⁸

But in the end, we must grapple with uncertainty as a feature of the landscape of practicing medicine. The lesson from the aforementioned case is not that uncertainty led to a patient's pregnancy loss. On the contrary, it was (misplaced) certainty that failed her. Patients who entrust themselves to our care are entitled to presume that we are knowledgeable, well-trained, and doing our best; they should not expect omniscience, nor should we expect it of ourselves. Among the challenges of practicing medicine is balancing "thorough enough" with "confident enough" and acting on that determination. We should be less concerned with zebras masquerading as horses or horses disguised as zebras than with our own fallibility. Modern tools, which include hard stops, checklists, algorithms, and digital alerts, may help reduce diagnostic errors, a utility that is likely to improve with time. But our patients look to us for something that (at least for the present) computers, algorithms, checklists, and hard stops cannot replace: our judgment. ■

REFERENCES

1. Hage PS, Arnouk IF, Zarou DM, Kim BH, Wehbeh HA. Laparoscopic management of ovarian ectopic pregnancy. *J Am Assoc Gynecol Laparosc* 1994;1:283–5.
2. Kadar N, DeVore G, Romero R. Discriminatory hCG zone: its use in the sonographic evaluation for ectopic pregnancy. *Obstet Gynecol* 1981;58:156–61.
3. Csapo AI, Pulkkinen MO, Wiest WG. Effects of luteectomy and progesterone replacement therapy in early pregnant patients. *Am J Obstet Gynecol* 1973;115:759–65.
4. Lewis M. *The undoing project: a friendship that changed our minds*. New York: WW Norton; 2016.
5. Kahneman D. *Thinking fast and slow*. New York: Farrar, Straus and Giroux; 2011.
6. Lambe KA, O'Reilly G, Kelly BD, Curristan S. Dual-process cognitive interventions to enhance diagnostic reasoning: a systematic review. *BMJ Qual Saf* 2016;25:808–20.
7. Balogh EP, Miller BT, Ball JR, editors. *Improving diagnosis in health care*. Washington (DC): National Academies of Sciences, Engineering, and Medicine; 2015.
8. Bartkowiak B, Snyder AM, Benjamin A, et al. Validating the electronic cardiac arrest risk triage (eCART) score for risk stratification of surgical inpatients in the postoperative setting: retrospective cohort study. *Ann Surg* 2019;269:1059–63.

ABSTRACT

Horses and zebras: probabilities, uncertainty, and cognitive bias in clinical diagnosis

Medical diagnosis is typically an iterative process guided by integration and synthesis of data into a model of disease. However, facts are not the only inputs into this process. A case of medical misdiagnosis is presented, in which systematic cognitive bias is considered to have played a role in generating error. Specific

cognitive biases are cited, and measures that can be taken to minimize their negative impact are reviewed.

Key words: cognitive bias, diagnostic process, medical error, ovarian ectopic pregnancy