Diagnostic Errors and Diagnostic Calibration

Adam S. Cifu, MD
University of Chicago,
Chicago, Illinois.

Even though medical errors have been a focus of research for decades, the subset of errors that occur during the diagnostic process have only recently been a major focus of quality improvement efforts.[1,2] The first and most important diagnostic test in every patient's care is the physician's assessment. This assessment dictates all further management and errors in the assessment reverberate throughout a patient's treatment course.

The study of diagnostic error is exceedingly challenging because traditional approaches to identifying errors are limited and successful remediation strategies are complex.[3] To a great extent, the reason that diagnostic errors are difficult to quantify and remediate is that these errors are based on the calibration of physicians, something that is difficult to measure and even more challenging to remediate.

Diagnostic calibration has been defined as the relationship between diagnostic accuracy and physician confidence in that accuracy.[4] In general, a physician with a low level of confidence in his or her level of diagnostic accuracy will overtest, whereas an overconfident physician will underevaluate patient concerns.

Although other diagnostic tests such as a serum chemistry assay or pleural fluid cell count are constantly titrated against standards to avoid error, physicians' diagnostic capabilities are not systematically

---

Diagnostic calibration has been defined as the relationship between diagnostic accuracy and physician confidence in that accuracy.

---

calibrated. Optimal diagnostic calibration depends on knowledge of epidemiology and diagnostic test characteristics, sensitivity to patient values, deliberate practice, and awareness of patient outcomes.

For example, a 50-year-old man with atypical chest pain schedules a visit with his internist. The physician interviews and examines the patient and finds no further clues regarding a diagnosis. Based on this assessment, the physician may use clinical decision rules to estimate that the patient's probability of having coronary disease is approximately 20%. The physician must know how or whether to further evaluate a patient with a low, but significant, probability of having a serious disease.

A competent physician might opt to reassure this patient by requesting an electrocardiogram, chest radiograph, stress test, or myriad laboratory tests. A well-calibrated physician will order tests associated with high specificities to diagnose diseases that are likely (but not

so likely that tests are unnecessary) and highly sensitive tests to exclude less likely, but important, diseases. There is no set number of positive or negative evaluations that define good calibration.

Clinical decision making begins with consideration of epidemiology and the test characteristics of the history and physical examination, quantitative factors that might appear to make this stage of the evaluation less dependent on physician calibration than later stages. In fact, because pretest probabilities and test characteristics are so imprecise, the course and outcomes of even this part of a patient evaluation are affected by calibration. The likelihood of a condition occurring in an individual patient with that patient's unique genetics, personal history, hobbies, and socioeconomics is never known.

Furthermore, findings from the history and physical examination are both imprecise and imperfect. For example, the intraobserver agreement for hearing an S3 gallop is remarkably low (κ = 0.14-0.37) and, when heard, is only of 13% sensitive (although quite specific) for the diagnosis of heart failure.[5,6]

What determines a physician's diagnostic calibration? Static influences affect the physician in more or less the same way from patient to patient and from day to day. One of these influences involves the physician's patient population. A physician with a relatively healthy patient population may be less likely to evaluate a given concern because, through years of experience with his or her population of patients, he or she has learned that most evaluations will be fruitless.

Each physician has a certain level of tolerance for risk and uncertainty. One physician may be comfortable following up a patient with chronic pain but no associated findings for years, whereas another physician can only be comfortable when the patient's symptom has been extensively evaluated.

Dynamic factors that influence a physician's diagnostic calibration include recent diagnostic experiences. This underlies the tendency to inflate the pretest probability for diagnoses that readily come to mind. The day-to-day mental state of a physician also affects decision making. It takes more energy to reason through a patient concern than it does to order a test. A physician is probably less likely to evaluate a patient with a low pretest probability on a day when he or she is well rested and focused than on a day when he or she is tired and overburdened.[7]

In addition, patients represent dynamic factors. Each patient has his or her own personal tolerance for risk. A 4% likelihood of cancer might mandate a biopsy for one patient, whereas this same likelihood

Corresponding Author: Adam S. Cifu, MD, University of Chicago, 5841 S Maryland Ave, MC 3051, Chicago, IL 60637 (adamcifu @uchicago.edu).

may be reassuring to another patient that no further workup is needed.

How do physicians and other clinicians reduce diagnostic errors when evaluations depend on an individual physician's diagnostic calibration? Current efforts focus on better defining the prevalence of diagnostic errors, their causes, and remediation strategies.[4] These are all crucial, but interventions should also acknowledge and address the role that physician calibration plays in diagnostic error such as considering standards for diagnostic testing and determining the best way of recalibrating the practice of physicians whose rates vary based on these standards.

Setting a diagnostic standard would mean determining an acceptable ratio of positive to negative evaluations for a given patient concern. A poorly calibrated physician might only receive positive results for a certain test (suggesting that he or she is missing diagnoses) or receive overwhelmingly negative results (suggesting he or she is overevaluating the patient concern). Setting a standard will be complex. Local norms could be used but this approach uses an average rather than a target and does not account for individual physicians' diagnostic abilities.

In addition, there are very few patient concerns that any one physician sees with great enough frequency to generate reliable data. Patient mix including disease prevalence and sociocultural factors that make office diagnoses more or less challenging also need to be considered in standard settings.

Because diagnostic calibration is the relationship between diagnostic accuracy and confidence in that accuracy, awareness of patient outcomes is critical. This is especially true because physicians tend to assume good patient outcomes for those with whom physicians lack follow-up. Feedback on patient outcomes needs to include a measure of patient satisfaction. A physician who orders fewer tests with good outcomes is well calibrated, but to what end is good calibration if that physician's patients are not satisfied with the care they are receiving?

The evaluation of every patient and every potential diagnostic error begins with a physician assessment, and the diagnostic assessment relies on the calibration of the physician supervising the assessment. Efforts to reduce diagnostic errors need to address physician calibration, most likely by providing feedback regarding test ordering and encouraging close attention to patient outcomes.

REFERENCES

1. National Academies of Sciences, Engineering, and Medicine. Improving diagnosis in health care. https://www.nap.edu/catalog/21794/improving-diagnosis-in-health-care. Accessed July 12, 2017.

2. Newman-Toker DE, Pronovost PJ. Diagnostic errors—the next frontier for patient safety. *JAMA.* 2009;301(10):1060-1062.

3. McGlynn EA, McDonald KM, Cassel CK. Measurement is essential for improving diagnosis and reducing diagnostic error: a report from the Institute of Medicine. *JAMA.* 2015;314(23):2501-2502.

4. Meyer AND, Payne VL, Meeks DW, Rao R, Singh H. Physicians' diagnostic accuracy, confidence, and resource requests: a vignette study. *JAMA Intern Med.* 2013;173(21):1952-1958.

5. Panju AA, Hemmelgarn BR, Guyatt GH, Simel DL. The rational clinical examination: is this patient having a myocardial infarction? *JAMA.* 1998;280(14):1256-1263.

6. Wang CS, FitzGerald JM, Schulzer M, Mak E, Ayas NT. Does this dyspneic patient in the emergency department have congestive heart failure? *JAMA.* 2005;294(15):1944-1956.

7. Linder JA, Doctor JN, Friedberg MW, et al. Time of day and the decision to prescribe antibiotics. *JAMA Intern Med.* 2014;174(12):2029-2031.