Issue Brief 9

Improved Diagnostic Accuracy Through Probability-Based Diagnosis

Agency for Healthcare Research and Quality

PATIENT SAFETY

This page intentionally left blank.

Issue Brief 9

Improved Diagnostic Accuracy Through Probability-Based Diagnosis

Prepared for:
Agency for Healthcare Research and Quality
5600 Fishers Lane
Rockville, MD 20857
www.ahrq.gov

Contract Number HHSP233201500022I/75P00119F37006

Prepared by:
Daniel J Morgan, M.D., M.S.
University of Maryland School of Medicine/VA Maryland Healthcare System

Ashley N.D. Meyer, Ph.D.
Center for Innovations in Quality, Effectiveness and Safety, Michael E. DeBakey VA Medical Center and Baylor College of Medicine

Deborah Korenstein, M.D.
Memorial Sloan Kettering Cancer Center

AHRQ Publication No. 22-0026-3-EF
September 2022

This project was funded under contract number HHSP233201500022I/75P00119F37006 from the Agency for Healthcare Research and Quality (AHRQ), U.S. Department of Health and Human Services. The authors are solely responsible for this document's contents, findings, and conclusions, which do not necessarily represent the views of AHRQ. Readers should not interpret any statement in this product as an official position of AHRQ or of the U.S. Department of Health and Human Services. None of the authors has any affiliation or financial involvement that conflicts with the material presented in this product.

Suggested citation: Morgan DJ, Meyer AND, Korenstein D. Improved diagnostic accuracy through probability-based diagnosis. Rockville, MD: Agency for Healthcare Research and Quality; September 2022. AHRQ Publication No. 22-0026-3-EF.

Introduction

Errors that occur during the diagnostic process can lead to missed or wrong diagnoses and can harm patients. Most patients will experience diagnostic errors in their lifetime.[1] Many diagnostic errors result from clinicians inadequately synthesizing clinical information,[2] such as weighing evidence and assigning proper probabilities to potential diagnoses.

Medical diagnosis is fundamentally based on probability.[3] Thus, more accurate execution of probability-based diagnosis is needed to reduce diagnostic errors related to poor information synthesis. For example, a negative stress ECG test in a patient at high risk for cardiac ischemia may be misinterpreted by someone with poor understanding of probability as definitive for the absence of disease, whereas it more often represents a false-negative. The result is a missed diagnosis and missed opportunity for early intervention.

Understanding probability and managing related uncertainty are vital for making accurate, timely diagnoses. Although central to clinical practice, these topics are not emphasized as critical skills or typically included in medical education curricula.[4,5] Despite longstanding public discussion around the need to improve diagnosis, probabilistic reasoning has not been emphasized as a core competency.[6] Better development and dissemination of evidence-based methods for training on probabilistic thinking could improve diagnostic accuracy and reduce errors.

This issue brief presents a framework that outlines the diagnostic process and highlights the role of probabilistic understanding at each step. The framework lays out the information needed throughout the diagnostic process to help clinicians make accurate probability assessments. It also proposes innovative methods to train on probabilistic reasoning skills for improving diagnostic decision making.

Fundamental Concepts for Understanding Probability

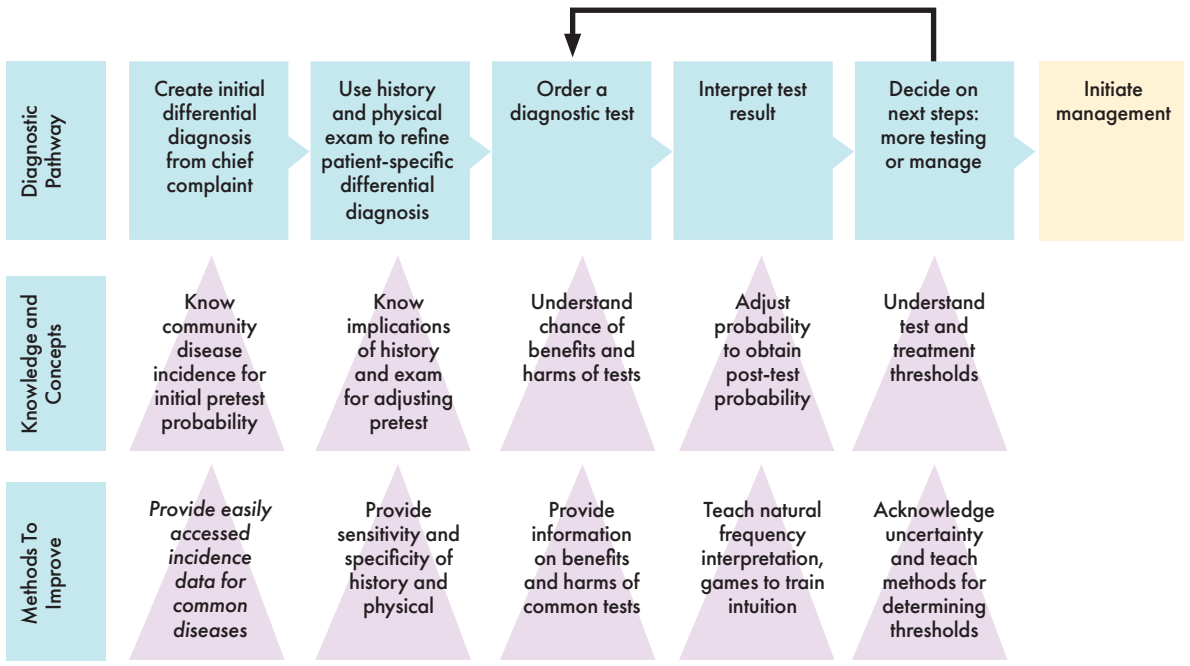
Embracing Uncertainty

"Uncertainty pervades the diagnostic process"[7] in the form of inherently imprecise estimates of probability, limits of scientific knowledge, and incomplete individual clinician knowledge.[8] But even with the most accurate probability estimates, high-quality evidence to inform decision making, and optimal clinician knowledge, the diagnostic process will always involve some uncertainty.

Few clinical decisions are made in the context of complete certainty. Beginning with the conceptualization of the diagnostic process created by the National Academies of Sciences, Engineering, and Medicine, we further mapped the process of diagnosis as it relates to probability.[1]

As illustrated in Figure 1, which depicts the framework we developed, every step of the process of diagnosis incorporates understanding of probabilities that are between 0 and 100 percent, with some uncertainty always present. For example, a patient presenting with acute cough, fever, and shortness of breath who has nonspecific findings on chest x ray is far from certain to have pneumonia, but treatment with antibiotics is likely appropriate. In the figure, steps in the diagnostic pathway are displayed in the boxes at the top and key points for understanding probability are shown in the boxes below.

Figure 1. Importance of probabilistic understanding in the diagnostic process



Source: Based on Stern SDC, Cifu AS, Altkorn D. Symptom to Diagnosis: An Evidence-Based Guide, Third Edition. New York, NY: McGraw-Hill; 2015. Chapter 1: Diagnostic Process.
https://accessmedicine.mhmedical.com/content.aspx?bookid=1088&sectionid=61696411.

The concept of uncertainty often provokes anxiety in clinicians, who may misconstrue it as a reflection of individual flaws, but uncertainty must be embraced to optimize diagnostic excellence. Clinicians who believe there is little uncertainty in medicine tend to overtest and overtreat patients.[9,10] Engaging with patients to explain and ideally quantify uncertainty can increase understanding, confidence, and accuracy of the diagnostic process.

Probabilistic reasoning that manages the inevitable uncertainty, including adjusting chance of disease and updating degree of uncertainty with new information during clinical care in an explicit fashion, may improve diagnosis. When diagnoses are expressed as possible and uncertainty is recognized, the chance of misdiagnosis and premature closure (failing to consider reasonable alternatives after an initial diagnostic impression) is reduced.[11]

Understanding the Importance of Language

Clear language reflects organized thinking. Use of nonspecific words can bias and confuse conversations around probability and muddy the diagnostic process. Simple language includes discussing chance or probability of disease or diseases, with numerical quantification.

For example, an 80-year-old man with hypertension and lower extremity edema who is worried about heart failure asks his primary care provider for a brain natriuretic peptide (BNP) test. The clinician might explain the low likelihood of heart failure, the impact of results of a BNP test, and the rationale for not performing the test by stating, "Someone like you has about a 2 percent chance of having heart failure. If we do a BNP test and it's abnormal, that chance of heart failure would rise to about 8 percent."

A focus on quantifying probability numerically is important given the inconsistency in interpretation of vague phrases such as "likely," "probably," and "maybe," which have been shown to mean different things to different people.[12] Adding to the imprecision, the medical community often further confounds understanding by relying on pseudo-probabilistic aphorisms such as "low threshold" and "diagnosis of exclusion."[13]

Biased language can influence clinical thinking and reasoning. For example, the outcome of medical decisions is often incorrectly phrased in terms of "risks" and "benefits" vs. the more neutral, transparent, and quantifiable comparison of "chance of harm" and "chance of benefit."[14]

Probability and the Diagnostic Pathway

Figure 1 depicts the steps in the diagnostic pathway and describes the probabilistic skills and information needed for each step. It shows the centrality of probabilistic understanding to making diagnoses at each of the following steps:

- **Step 1: Chief complaint and initial differential diagnosis.** The diagnostic pathway begins with a chief complaint from a patient. The clinician creates an initial differential diagnosis based on *knowledge of disease incidence in the population of interest (i.e., the population the patient is from)*. This estimate serves as the initial "pretest probability."
- **Step 2: Adjustment based on history and physical exam.** The clinician performs a history and physical examination, each element of which informs *adjustment of the probabilities of diseases in the differential diagnosis*, ending in a patient-specific differential diagnosis. This step requires knowledge of the impact of each history and physical exam finding on probability (e.g., test accuracy), expressed as *sensitivity and specificity or likelihood ratios*.
- **Step 3: Selection of diagnostic tests.** The clinician then decides to perform a particular diagnostic test (or set of tests) to explore likely diagnoses. The chosen tests may relate to the most likely diagnosis or the most concerning diagnosis, depending on the clinical scenario. Deciding to order a test requires understanding of *the probability that the patient may benefit or be harmed by getting the test*.

  While many clinicians do not frame diagnostic testing in terms of patient benefit and harms, tests, like all other health services, will either help patients or harm them. For example, a test may identify a diagnosis for which treatment improves outcomes or it may expose a patient to toxic substances, inconvenience, or unnecessary care for which harms outweigh benefits. These benefits and harms vary widely in magnitude.

- **Step 4: Test interpretation.** Once a test is performed, the clinician must *interpret the results in the context of the pretest probability to arrive at a posttest probability*. This step requires understanding Bayes Theorem, which integrates measures of test accuracy into the pretest probability and requires rejecting the notion that test results are definitive.[15]

  While explicit calculations using Bayes Theorem may not be feasible during clinical practice, conceptual understanding of Bayes Theorem informs clinical thinking. For example, a 40-year-old woman with no cardiac risk factors and nonspecific chest pain has an abnormal exercise stress test. She remains *unlikely* to have coronary artery disease, because her pretest probability was so low. If multiple tests are done, the results of each should be considered when calculating the ultimate posttest probability.

- **Step 5: Final diagnosis or further testing.** Finally, the clinician must decide when diagnostic closure is achieved to complete the diagnostic phase. This step involves determining whether a diagnosis is established, i.e., considering *whether the posttest probability is high enough (or the uncertainty is low enough) to begin management*. The disease probability at which that threshold is crossed (the "treatment threshold") will vary based on characteristics of the disease and its treatment, as well as clinician and patient risk tolerance.[16] If the likelihood of disease is lower than the threshold, further testing may be appropriate.

Correct diagnosis, then, relies on accurate estimates of pretest probability and understanding the influence of positive and negative tests on that probability. However, clinicians generally overestimate the chance a patient has disease under consideration, both before and after testing.[17] This tendency likely leads to misdiagnosis of conditions from false-positive test results and subsequent missed diagnoses that are truly causing symptoms, with potential for patient harm.

Reasons To Refocus Training on Probability

While the need to weigh probabilities is widely accepted as foundational to the diagnostic process and medical students may receive instruction in test accuracy, clinically integrated training in probabilistic diagnostic reasoning is lacking.[18] Learning about probability has mostly been limited to memorizing definitions of sensitivity and specificity and calculating them from studies of test accuracy using 2X2 tables.

Errors in estimating probability of disease may arise from this approach[19] as mathematical calculations are difficult to apply to clinical medicine.[20] Indeed, most teaching about test interpretation and probability happens in the preclinical years of medical school, suggesting that it is separate from the approach to testing used in daily clinical practice.

Further, in clinical practice, most medical decisions are made rapidly and intuitively,[21] so it is critical for clinicians to make accurate "gestalt" estimates of pretest probabilities of common disorders and intuitively adjust pretest probabilities based on test results.[22] These estimates must be updated with each subsequent test. At the same time, for less common presenting complaints or syndromes, for which clinicians have little intuition, they must use quantitative estimates to make decisions.

Currently, little evidence is available to inform approaches to teaching diagnostic reasoning and the most common discussions of diagnosis for trainees occur in the context of generating differential diagnoses.[23] These discussions do not emphasize probabilistic understanding. In fact, they may undermine intuitive understanding of prevalence and probabilistic reasoning by rewarding learners for suggesting rare diseases with extremely small likelihood.

Methods To Teach and Inform Probability for Diagnostic Decision Making

To appropriately incorporate probabilistic thinking into the diagnostic process, clinicians need to both have ready access to the variety of critical quantitative data that is currently difficult to obtain and intuitively understand probability. Figure 1 shows the probabilistic information needed at each step in the diagnostic process and current needs to fill the gap.

- **At Step 1: Developing initial differential diagnosis:** *Provide easily accessible incidence data for common diseases*. When estimating an initial differential diagnosis based on the chief complaint, clinicians need ready access to incidence data in the local community. This information may be challenging to obtain but it is possible. For example, information about local rates of COVID-19 can be easily accessed online (e.g., through the Centers for Disease Control and Prevention's COVID Data Tracker) and are useful to facilitate test interpretation.
- **At Step 2: Adjustment based on history and physical:** *Provide sensitivity and specificity of history and physical examination data*. When adjusting the initial differential based on elements of the history and physical, clinicians need evidence of the sensitivity and specificity of clinical features and physical examination findings, which are largely unavailable. In addition, they need to understand how to adjust probability based on test results, whether those tests are physical examination maneuvers or blood or imaging tests. Such information is difficult to find but exists most prominently in JAMA's Rational Clinical Examination series at https://jamanetwork.com/collections/6257/the-rational-clinical-examination.
- **At Step 3: Selection of a diagnostic test:** *Provide information on probability of potential benefits and harms of common tests in absolute terms that is readily available at the point of care*. While few clinicians currently conceptualize tests through a lens of benefits and harms, the approach to screening tests is an illustrative exception. For example, using currently available decision aids, we can frame decisions around PSA tests as a balance between potential benefits (e.g., reduction in prostate cancer deaths or advanced disease) and harms (e.g., urinary incontinence and erectile dysfunction). Similar logic could be applied to testing more generally with better access to information.
- **At Step 4: Test interpretation:** *Teach about test accuracy using natural frequency interpretation via games or other novel methods*. Perhaps most importantly, clinicians need better intuitive understanding of probability. Teaching probability can be most effective when students are exposed to natural frequency figures.[24]

  Instead of tables with probability calculations, approaches can be grounded in patient populations. For example, students may be asked to consider 100 identical patients, then divide them by the percentage likely to have (representing pretest probability) or not have disease. The number of positive tests that would occur in patients with disease and those without disease could be estimated based on incidence and sensitivity and specificity, respectively.[25] These estimates can be depicted using graphic images showing grids or icons representing risks out of 100 or 1,000 people; such illustrations have been found to work better than other depictions in those with less training.[26]

  To our knowledge, these techniques have not been widely studied or adopted in medical training. In addition, repeated practice with clear, actionable feedback is a classic and effective learning approach that can help clinicians develop an intuitive sense of probability.[27, 28]

- **At Step 5: Final diagnosis or further testing or consideration of other diagnoses:** *Acknowledge uncertainty and teach methods for determining thresholds*. Diagnostic error in clinical medicine exists in part because of the inherent uncertainty that stems from the great diversity in patient symptoms and findings and the lack of clarity around many diagnoses. Clinicians confronting this uncertainty seldom receive feedback about the assumptions that underpin their diagnoses, which can reinforce faulty reasoning.

  Educators themselves may need instruction in managing uncertainty.[29] Probability thresholds for testing and treating may vary by individual and geographic region, but appropriate ranges can be estimated based on survey studies.[30] Probabilistic treatment thresholds should be included in all clinical discussions of plans for diagnostic testing. This information can reinforce the importance of probability and provide feedback that can fine-tune learners' sense of appropriate thresholds for testing or treatment in different contexts.

Novel delivery of classic learning approaches can serve as an effective method for teaching these skills. For example, games with a primary educational goal, known as "serious games," use repetitive, rapid decision making with immediate feedback to train skills.[31] They have been widely used to improve skill in areas such as chess and gambling. More recently, these games have moved to medicine, where successful applications have included patient care simulations in emergency surgical settings. Such games can be more efficient than standard problem-based learning and may be superior for training intuitive skill (vs. improving knowledge).[32]

Most games have focused on discrete lessons related to individual cases, with few trying to develop a general skill. More recently, games have targeted heuristics to change thinking processes inherent in clinical medicine, suggesting broad future application.[33, 34] One such game resulted in durable improvement in appropriate triage of trauma patients in an emergency department.[33] Serious games have potential to facilitate achievement of diagnostic excellence in medicine by motivating repetition and feedback and could be used both during training and by practicing clinicians, ideally for continuing education credit.

Future Vision for Probabilistic Diagnostic Decisions

This framework illustrates how future clinicians can be equipped to make more accurate diagnoses and reduce error through attention to probabilistic diagnosis. Once internalized, this approach is largely intuitive and uses information provided at the point of care, and it would not add substantial time or cognitive load to clinical encounters. Improving diagnostic accuracy will require better clinician skills and tools, including the following achievable steps:

- Focus on embracing uncertainty as a core educational principle.[29,35] Medical school and other health professional courses on pathophysiology and history and physical examination should acknowledge and quantify the high degree of uncertainty embedded in all clinical care. Such acknowledgment could be easily integrated into current medical school curricula. Curricula for other clinicians such as nurse practitioners currently include very little content on probabilistic thinking and uncertainty; integration in these settings will be critical but may require more substantial change.

- Emphasize quantifying probability throughout medical education. During clinical training, students and residents should be asked to quantify the likelihood of different diagnoses, expressing it as a probability or range of probabilities. This exercise will highlight the incidence of common diseases in various populations, train intuition, and trigger discussions of determinants of those probabilities to inform more formal probabilistic reasoning.
- Enable clinician access to better data on disease incidence. This goal could partly be accomplished by incorporating links to available data, such as the CDC COVID Data Tracker, into electronic health records (EHRs). Ideally, the approach would use novel EHR-based tools that could refine estimates using clinical characteristics (such as disease calculators on testingwisely.com). It would also require development of better evidence on pretest probability based on epidemiologic studies that report both disease incidence in the population and the nature and frequency of presentations in clinical settings.
- Provide data about test performance at the point of care, ideally through integration with EHRs. Various methods are available to structure such information delivery; doing so at the time of test ordering would be optimal.

Ultimately, better clinician management of probability will lead to better management of patients and fewer diagnostic errors. Clinicians must be better educated to accurately estimate disease probabilities in the general population and in individual patients and to adjust those probabilities in response to test results. Education must train probabilistic intuition, empower more deliberative probability adjustment, and provide needed tools at the point of care. Through understanding probability, clinicians can improve the diagnostic process and optimize patient safety.

References

1. Committee on Diagnostic Error in Health Care; Board on Health Care Services; Institute of Medicine; National Academies of Sciences, Engineering, and Medicine. Improving Diagnosis in Health Care. Balogh EP, Miller BT, Ball JR, eds. Washington, DC: National Academies Press; 2015. https://www.ncbi.nlm.nih.gov/books/ NBK338596/. Accessed August 31, 2022.
2. Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. Arch Intern Med. 2005;165(13):1493-9. https://pubmed.ncbi.nlm.nih.gov/16009864/. Accessed August 31, 2022.
3. Weatherall M. Information provided by diagnostic and screening tests: improving probabilities. Postgrad Med J. 2018;94(1110):230-5. https://pubmed.ncbi.nlm.nih.gov/29133377/. Accessed August 31, 2022.
4. Graber ML, Holmboe E, Stanley J, Danielson J, Schoenbaum S, Olson APJ. A call to action: next steps to advance diagnosis education in the health professions. Diagnosis (Berl). 2021;9(2):166-75. https://pubmed.ncbi.nlm.nih.gov/34881533/. Accessed August 31, 2022.
5. Pauker SG, Kassirer JP. The threshold approach to clinical decision making. N Engl J Med. 1980; 302(20):1109-17. https://www.nejm.org/doi/10.1056/NEJM198005153022003. Accessed August 31, 2022.
6. Olson A, Rencic J, Cosby K, Rusz D, Papa F, Croskerry P, Zierler B, Harkless G, Giuliano MA, Schoenbaum S, Colford C, Cahill M, Gerstner L, Grice GR, Graber ML. Competencies for improving diagnosis: an interprofessional framework for education and training in health care. Diagnosis (Berl). 2019;6(4):335-41. https://pubmed.ncbi.nlm.nih.gov/31271549/. Accessed August 31, 2022.
7. Dahm MR, Crock C. Understanding and communicating uncertainty in achieving diagnostic excellence. JAMA. 2022;327(12):1127-8. DOI: 10.1001/jama.2022.2141. Accessed August 31, 2022.
8. Meyer AND, Giardina TD, Khawaja L, Singh H. Patient and clinician experiences of uncertainty in the diagnostic process: current understanding and future directions. Patient Educ Couns. 2021;104(11):2606-15. https://pubmed.ncbi.nlm.nih.gov/34312032/. Accessed August 31, 2022.
9. Korenstein, D, Scherer LD, Foy A, Pineles L, Lydecker AD, Owczarzak J, Magder L, Brown JP, Pfeiffer CD, Terndrup C, Leykum L, Stevens D, Feldstein DA, Weisenberg SA, Baghdadi JD, Morgan DJ. Clinician attitudes and beliefs associated with more aggressive diagnostic testing. Am J Med. 2022 Jul;135(7):e182-e193. https://pubmed.ncbi.nlm.nih.gov/35307357/. Accessed August 31, 2022.
10. Baghdadi JD, Korenstein D, Pineles L, Scherer LD, Lydecker AD, Magder L, Stevens DN, Morgan DJ. Exploration of primary care clinician attitudes and cognitive characteristics associated with prescribing antibiotics for asymptomatic bacteriuria. JAMA Netw Open. 2022 May 2;5(5):e2214268. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9142875/. Accessed August 31, 2022.
11. McSherry D. Avoiding premature closure in sequential diagnosis. Artif Intell Med. 1997;10(3):269-83. https://pubmed.ncbi.nlm.nih.gov/9232189/. Accessed August 31, 2022.
12. Ott DE. Words representing numeric probabilities in medical writing are ambiguous and misinterpreted. JSLS. 2021;25(3):e2021.00034. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8325478/. Accessed August 31, 2022.
13. Bergl PA. Pseudo-probabilistic aphorisms. J Grad Med Educ. 2018;10(6):709-10. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6314368/. Accessed August 31, 2022.

14. Morgan DJ, Scherer LD, Korenstein D. Improving physician communication about treatment decisions: reconsideration of "risks vs benefits." JAMA. 2020;324(10):937-8.
15. Bours MJ. Bayes' rule in diagnosis. J Clin Epidemiol. 2021;131:158-60. https://pubmed.ncbi.nlm.nih.gov/33741123/. Accessed August 31, 2022.
16. Djulbegovic B, van den Ende J, Hamm RM, Mayrhofer T, Hozo I, Pauker SG; International Threshold Working Group (ITWG). When is rational to order a diagnostic test, or prescribe treatment: the threshold model as an explanation of practice variation. Eur J Clin Invest. 2015;45(5):485-93. https://pubmed.ncbi.nlm.nih.gov/25675907/. Accessed August 31, 2022.
17. Morgan DJ, Pineles L, Owczarzak J, Magder L, Scherer L, Brown JP, Pfeiffer C, Terndrup C, Ley-kum L, Feldstein D, Foy A, Stevens D, Koch C, Masnick M, Weisenberg S, Korenstein D. Accuracy of practitioner estimates of probability of diagnosis before and after testing. JAMA Intern Med. 2021;181(6):747-55. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8022260/. Accessed August 31, 2022.
18. Cooper N, Bartlett M, Gay S, Hammond A, Lillicrap M, Matthan J, Singh M; UK Clinical Reasoning in Medical Education (CReME) consensus statement group. Consensus statement on the content of clinical reasoning curricula in undergraduate medical education. Med Teach. 2021;43(2):152-9. https://pubmed.ncbi.nlm.nih.gov/33205693/. Accessed August 31, 2022.
19. Noguchi Y, Matsui K, Imura H, Kiyota M, Fukui T. Quantitative evaluation of the diagnostic thinking process in medical students. J Gen Intern Med. 2002;17(11):839-44. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1495132/. Accessed August 31, 2022.
20. Brighton H, Gigerenzer G. Homo heuristicus: less-is-more effects in adaptive cognition. Malays J Med Sci. 2012;19(4):6-16. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3629675/. Accessed August 31, 2022.
21. Djulbegovic B, Hozo I, Beckstead J, Tsalatsanis A, Pauker SG. Dual processing model of medical decision-making. BMC Med Inform Decis Mak. 2012 Sep 3;12:94. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3471048/. Accessed August 31, 2022.
22. Dale AP, Marchello C, Ebell MH. Clinical gestalt to diagnose pneumonia, sinusitis, and pharyngitis: a meta-analysis. Br J Gen Pract. 2019;69(684):e444-e453. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6582453/. Accessed August 31, 2022.
23. Xu H, Ang BWG, Soh JY, Ponnamperuma GG. Methods to improve diagnostic reasoning in undergraduate medical education in the clinical setting: a systematic review. J Gen Intern Med. 2021;36(9):2745-54. https://pubmed.ncbi.nlm.nih.gov/34159542/. Accessed August 31, 2022.
24. Hoffrage U, Gigerenzer G. Using natural frequencies to improve diagnostic inferences. Acad Med. 1998;73(5):538-40. https://pubmed.ncbi.nlm.nih.gov/9609869/. Accessed August 31, 2022.
25. Johnson KM. Using Bayes' rule in diagnostic testing: a graphical explanation. Diagnosis (Berl). 2017;4(3):159-67. https://pubmed.ncbi.nlm.nih.gov/29536931/. Accessed August 31, 2022.
26. Zipkin DA, Umscheid CA, Keating NL, Allen E, Aung K, Beyth R, Kaatz S, Mann DM, Sussman JB, Korenstein D, Schardt C, Nagi A, Sloane R, Feldstein DA. Evidence-based risk communication: a systematic review. Ann Intern Med. 2014;161(4):270-80. https://www.acpjournals.org/doi/10.7326/M14-0295. Accessed August 31, 2022.

27. Hysong S.J. Meta-analysis: audit and feedback features impact effectiveness on care quality. Med Care. 2009;47(3):356-63. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4170834/. Accessed August 31, 2022.

28. Ebbinghaus EE. Memory: A Contribution to Experimental Psychology. Ruger HA, Bussenius CE, translators. New York, NY: Dover; 1964.

29. Moffett J, Armitage-Chan E, Hammond J, Kelly S, Pawlikowska T. "It's okay to not know …" a qualitative exploration of faculty approaches to working with uncertainty. BMC Med Educ. 2022;22(1):135. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8887020/. Accessed August 31, 2022.

30. Ebell MH, Locatelli I, Senn N. A novel approach to the determination of clinical decision thresholds. Evid Based Med. 2015;20(2):41-7. https://pubmed.ncbi.nlm.nih.gov/25736042/. Accessed August 31, 2022.

31. Girard C, Ecalle J, Magnan A. Serious games as new educational tools: how effective are they? A meta-analysis of recent studies. J Comput Assist Learn. 2012;29(3):207-19. https://doi.org/10.1111/j.1365-2729.2012.00489.x. Accessed August 31, 2022.

32. Graafland M, Dankbaar M, Mert A, Lagro J, De Wit-Zuurendonk L, Schuit S, Schaafstal A, Schi-jven M. How to systematically assess serious games applied to health care. JMIR Serious Games. 2014;2(2):e11. https://games.jmir.org/2014/2/e11/. Accessed August 31, 2022.

33. Mohan D, Farris C, Fischhoff B, Rosengart MR, Angus DC, Yealy DM, Wallace DJ, Barnato AE. Efficacy of educational video game versus traditional educational apps at improving physician decision making in trauma triage: randomized controlled trial. BMJ. 2017 Dec 12.359:j5416. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5725983/. Accessed August 31, 2022.

34. Mohan D, Schell J, Angus DC. Not thinking clearly? Play a game, seriously! JAMA. 2016;316(18):1867-8. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5654519/. Accessed August 31, 2022.

35. Stephens GC, Rees CE, Lazarus MD. Exploring the impact of education on preclinical medical students' tolerance of uncertainty: a qualitative longitudinal study. Adv Health Sci Educ Theory Pract. 2021;26(1):53-77. https://pubmed.ncbi.nlm.nih.gov/32378150/. Accessed August 31, 2022.

AHRQ Pub. No. 22-0026-3-EF
September 2022