EDITORIALS

Five Uneasy Pieces About Pre-test Probability

In this issue, Noguchi et al. report on their study of Japanese medical students' abilities to estimate pre-test probabilities, test characteristics, and post-test probabilities for written cases representing patients with differing likelihoods of coronary heart disease, compared with reference standards from the published literature.¹ They found among other things that students made inaccurate estimates of pre-test probability, particularly in scenarios with lower likelihoods of disease. These data give us pause to confront the uneasy predicament we face with pre-test probability, expressed in 5 questions.

1. How, and How Accurately, Do Clinicians Estimate Pre-test Probability?

What we have been taught about this seems so simple—as we see patients, we use our expertise to recognize their clinical problems, recall prior patients with the same problems, and use the diagnoses made for prior patients to estimate pre-test probabilities for each new patient.^{2,3} Yet when clinicians are asked to estimate the probability of disease in experimental studies, they have generally done quite poorly, in that their estimates have shown low accuracy and wide variations.^{4–7} Although further research is needed, the data of Noguchi et al., when taken together with prior studies, suggest that both beginning and experienced clinicians often make inaccurate estimates of pre-test probability.

2. What Can Explain Clinicians' Difficulties in Estimating Pre-test Probability?

Some of the explanation can be understood as the limitations inherent in the 2 sources traditionally taught for pre-test probability, clinical experience and population prevalence. Commentators have interpreted the above-mentioned research findings as simply illustrations of flaws in general human reasoning that can distort probabilities, flaws to which we are not immune by being clinicians.^{8–11} We tend to remember recent or remarkable patients, who represent numerators, without retrieving the other, less memorable patients who fill the denominator. Thus, our unaided memories of prior cases can lead us to miscalculate the fractions involved in estimating probability.

Population prevalence is also touted as useful when estimating pre-test probability. While this might be true for making screening decisions, prevalence is much less useful for the clinical diagnosis of ill persons, for 2 reasons. The first is pragmatic—it remains nearly impos-

sible to get integrated, “real time” access to current, accurate prevalence figures for each disorder being considered in a given patient's illness. The second reason is that population prevalence represents a fraction with the wrong denominator, that is, the whole population, including both sick and well persons.² Instead, when estimating pre-test probability for a sick person with a specific clinical problem, say syncope or fever of unknown origin, we'd rather know the likelihood of specific diseases among patients with that same clinical problem.^{2,12,13}

With the limitations of these 2 sources in mind, it shouldn't surprise us that clinicians of all levels of experience have some difficulty in estimating pre-test probability accurately. In addition to these “input” limitations, clinicians have also shown difficulty in revising probabilities after new information, such as test results, arrives.^{14,15}

3. What Are the Consequences of Inaccurate Estimates of Pre-test Probability?

Standard teaching of the Bayesian approach to diagnosis usually includes the admonition that the proper selection and interpretation of diagnostic tests begins with an estimate of the pre-test probability of the target disorder.³ Since inaccurate pre-test probabilities will translate into inaccurate post-test probabilities, we might expect the consequences of this inaccuracy to include poor test selection, poor interpretation of results, and ultimately diagnostic error. Although too little is known about how often these mistakes occur, inaccurate probability estimation is recognized as a substantial contributor to diagnostic error.^{14–16}

4. Could Evidence from Clinical Care Research Help Us Estimate Pre-test Probability?

Two different forms of clinical care research can be used to guide our estimates of pre-test probability. First, there are direct studies of disease probability, wherein investigators assemble a large and representative cohort of patients with a defined clinical problem, carry out careful diagnostic evaluations, apply explicit diagnostic criteria, and report the frequency of the underlying disorders that caused the patients' illnesses.¹⁷ Widely known examples include studies of patients with syncope^{18,19} and fever of unknown origin.^{20,21} If our appraisal suggests they are valid and applicable to our practices, we can use the disease probability results as starting estimates of pre-test

probability for our own patients, and then adjust the probabilities after considering characteristics of our patients or our practices.^{12,22}

Second, there are studies of clinical decision rules, wherein investigators may assemble similar cohorts of patients suspected to have the target disorder, subject them to reference standard tests, and report the frequency of diagnosis of the target disorder in subgroups with differing clinical features.²³ A widely known example of this type of evidence is a decision rule for pulmonary embolus.²⁴ If our appraisal suggests they are valid and applicable to our practices, we can match our patients' clinical features with those of the decision rule and use the results as estimates of pre-test probability for the target disorder.

5. When Should We Use Evidence to Guide Our Estimates of Pre-test Probability?

Little research addresses this question, so the following suggestions are tentative. First, beginning clinicians will find such evidence useful to *educate* their estimates of disease probability, while experienced clinicians may find it more useful to *recalibrate* themselves after an unusual case.²² Second, clinicians may find evidence about disease probability more useful for patients with clinical problems they see less often than for those with problems they see every day. Third, clinicians may also find this evidence more useful for the lower likelihood conditions they hope to exclude with testing (their "active alternatives") than for the higher likelihood condition they plan to confirm with testing (their "leading hypothesis").¹⁷

Much work remains to be done, both in making more research evidence about disease probability, and in the development of better methods to find it, appraise it, synthesize it, and make it available to clinicians at the point of care. While this field evolves, we can learn better how to express our diagnostic uncertainty in probability terms and how to use evidence to guide our estimates. As for most mathematical thinking we're not born with, learning this discipline may not be easy at first, yet if we stick with it our competence and our confidence can grow. In doing so, we should become better at selecting differential diagnoses, choosing and interpreting diagnostic tests, and reducing diagnostic error. If our current uneasiness impels us to take these steps, our patients and our learners will reap the rewards.—**W. SCOTT RICHARDSON, MD**, Department of Internal Medicine, Wright State University School of Medicine and Three Owl Learning Institute, Dayton, Ohio.

REFERENCES

1. 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;11:839-44.

2. Fletcher RH, Fletcher SW, Wagner EH. *Clinical Epidemiology: the Essentials*, Third ed. Baltimore, Md: Williams and Wilkins; 1996:43-74.
3. Black ER, Bordley DR, Tape TG, Panzer RJ, eds. *Diagnostic Strategies for Common Medical Problems*, Second ed. Philadelphia, Pa: American College of Physicians; 1999.
4. Dolan JG, Bordley DR, Mushlin AI. An evaluation of clinicians' subjective prior probability estimates. *Med Decis Making*. 1986; 6:216-23.
5. Bobbio M, Detrano R, Shandling AH, et al. Clinical assessment of the probability of coronary artery disease: judgmental bias from personal knowledge. *Med Decis Making*. 1992;12:197-203.
6. Bobbio M, Fubini A, Detrano R et al. Diagnostic accuracy of predicting coronary artery disease related to patients' characteristics. *J Clin Epidemiol*. 1994;47:389-95.
7. Lyman GH, Balducci L. The effect of changing disease risk on clinical reasoning. *J Gen Intern Med*. 1994;9:488-95.
8. Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. *Science*. 1974;185:1124-31.
9. Dawson NV, Arkes HR. Systematic errors in medical decision-making: judgment limitations. *J Gen Intern Med*. 1987;2:183-7.
10. Bornstein BH, Emler AC. Rationality in medical decision making: a review of the literature on doctors' decision-making biases. *J Eval Clin Pract*. 2001;7:97-107.
11. Redelmeier DA, Ferris LE, Tu JV, Hux JE, Schull MJ. Problems for clinical judgment: introducing cognitive psychology as one more basic science. *Can Med Assoc J*. 2001;164:358-60.
12. Richardson WS. Where do pretest probabilities come from [Editorial]? *Evidence-Based Medicine*. 1999;4:68-69.
13. Hilden J. Prevalence-free utility-respecting indices of diagnostic power do not exist. *Stat Med*. 2000;19:431-40.
14. Kassirer JP, Kopelman RI. Cognitive errors in diagnosis: instantiation, classification and consequences. *Am J Med*. 1989;86: 433-41.
15. Elstein AS, Schwartz A. Clinical problem solving and diagnostic decision making: a selective review of the cognitive research literature. In: Knottnerus JA, ed. *The Evidence Base of Clinical Diagnosis*. London: BMJ Books; 2002:179-95.
16. Bordage G. Why did I miss the diagnosis? Some cognitive explanations and educational implications. *Acad Med*. 1999; 74(10S):S138-43.
17. Richardson WS, Wilson MC, Guyatt GH, Cook DJ, Nishikawa J, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature. XV. How to use an article about disease probability for differential diagnosis. *JAMA*. 1999;281: 1214-9.
18. Kapoor WN. Evaluation and outcome of patients with syncope. *Medicine (Baltimore)*. 1990;69:160-75.
19. Getchell WS, Larsen GC, Morris CD, McNulty JH. Epidemiology of syncope in hospitalized patients. *J Gen Intern Med*. 1999;11: 677-87.
20. Petersdorf RG, Beeson PB. Fever of unexplained origin: report on 100 cases. *Medicine (Baltimore)*. 1961;40:1-30.
21. Larson EB, Featherstone HJ, Petersdorf RG. Fever of undetermined origin: diagnosis and follow-up of 105 cases. 1970 - 1980. *Medicine (Baltimore)*. 1982;61:269-92.
22. Richardson WS, Glasziou P, Polashenski WA, Wilson MC. A new arrival: evidence about differential diagnosis [Editorial]. *ACP J Club*. 2000;133:A11-2.
23. McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature. XXII. How to use articles about clinical decision rules. *JAMA*. 2000;284:79-84.
24. Wells PS, Ginsberg JS, Anderson DR, et al. Use of a clinical model for safe management of patients with suspected pulmonary embolism. *Ann Intern Med*. 1998;129:997-1005.