Estimating disease likelihood: a case of rubbery figures

In diagnosis and prognosis, we should avoid intuitive "guesstimates" and seek a validated numerical aid

ONE OF THE AXIOMS of clinical practice is that, in medicine, there are few, if any, certainties. When assessing the likelihood of a specific disease in a particular patient, or the chance of a future adverse event in a patient with known disease, clinicians are estimating probabilities or risk. These estimates derive from a clinical gestalt — the process of interpreting findings from history, examination and simple investigations (diagnosis), or of disease-specific correlates of complications or death (prognosis).

Clinicians use these estimates of probability or risk to decide whether they should intervene immediately, particularly if effective treatments are available. Alternatively, if the disease likelihood is low, or treatments toxic or only marginally effective, these estimates are used to decide whether to defer treatment and either observe expectantly or conduct more sophisticated tests whose results may substantially alter pre-test likelihood estimates.[1] If the estimate is too high, patients may incur unnecessary treatments or confirmatory investigations, or, if the estimate is too low, they may suffer the consequences of delayed intervention.

Thus, a fair bit is riding on how accurately we can judge the likelihood of current or future disease. Available research suggests that, for various reasons, we are not that good at it.[2-4] Common pitfalls include:

- framing a clinical problem in a way that may exaggerate risk;
- overweighting or underweighting certain clinical features;

- erroneously extrapolating past, vividly recalled cases to current patients; or
- manipulating risk subliminally to better fit with a preferred course of action (or inaction).

Overall, most of us, not surprisingly, are risk averse and will commit to action to avoid personal regret at witnessing an unfavourable but possibly preventable event, even if our perception of risk of such an occurrence seems low.[5]

In this issue of the Journal, Attia and colleagues (*page 449*) evaluate the extent to which clinicians' estimates of probability or risk for commonly encountered case scenarios vary from the "correct" estimate, and which clinician-related factors may influence such variation.[6] They distributed three hypothetical case scenarios to groups of general practitioners and physicians in Australia and the United Kingdom, and compared respondents' estimated probabilities of angina (in a patient with chest pain), deep vein thrombosis (DVT) (in a patient with a swollen leg), and future stroke (in a patient with chronic atrial fibrillation) with the "correct" estimates derived from statistically validated clinical-decision rules.

Two cautions come to mind:

- were the clinicians given sufficient information on which to base a reasoned judgement (keeping in mind that they could not examine the patients); and
- how accurate was the rule-based estimate as the reference standard?

One could argue that, in the chest-pain scenario, few experienced clinicians would be comfortable estimating the likelihood of angina simply on being told of a 65-year-old man presenting with exertional chest pain, without more detail about the character of the pain, the existence of coronary risk factors, and any signs of vascular disease seen on physical examination. The decision rule applied to the same case is also suspect, as it includes, for example, rapid relief with nitroglycerine as being positively predictive of angina, which recent evidence would challenge.[7] In the other two scenarios, the clinical details provided were more complete, and the decision rules more robust.

Another concern is that the "correct" estimate was stated as a single percentage, which clinicians were expected, perhaps unfairly, to closely approximate. This ignored the fact that, in developing the rule, the "correct" estimate is actually a mean within a range of observed frequencies, all of which would probably lead to the same clinical action.

On the positive side, the strengths of the study were its large, representative samples of clinicians, use of three different scenarios, and use of logistic regression to identify clinician-specific predictors of accuracy.

Setting aside methodological limitations, how did the respondents fare? Only slightly more than half of the whole group were within 20 percentage points of the "correct" probability estimate for the angina and stroke scenarios, and less than one in 10 achieved a similar result with the DVT scenario. In keeping with my earlier comments, most respondents overestimated rather than underestimated the risk, with estimates spread over a huge range, from 10% to 100% at least, for all cases. There was a notable lack of association between accuracy and experience as measured by age, years of practice, or field of specialty, with GPs performing as well as physicians. Unfortunately, the study by Attia et al did not have the power to determine whether graduating from a medical course that used problem-based learning — with emphasis on evidence appraisal — predisposed to better performance.

The implications of this study and others are several. First, all clinicians, irrespective of experience, appear to have problems quantifying probability or risk of disease, and, while there may be exceptions, this difficulty is independent of the clinical circumstances. Consequently, we should avoid intuitive "guesstimates" and seek instead a validated decision-rule, scoring scheme or other numerical aid that gets us closer to the mark. Fortunately, an increasing number of such tools are becoming available[8] and in a form compatible with hand-held computers. Second, if we are to choose the best rules and use them appropriately, we need to understand how such rules should be constructed and tested.[9] Third, we may need to "unlearn" some of our cherished clinical "rules of thumb" if evidence arises that questions their validity.[10] Finally, we should advocate for more research into decision aids that will help us to more accurately estimate and communicate likelihood of disease in individual patients. The results of such efforts should facilitate a more rational use of investigations and treatments and lead to better patient outcomes.

Ian A Scott

Director of Internal Medicine, Princess Alexandra Hospital
Associate Professor of Medicine, University of Queensland, Brisbane, QLD
ian_scott@health.qld.gov.au

1. Pauker SG, Kassirer JP. The threshold approach to clinical decision making. *N Engl J Med* 1980; 302: 1109-1117.
2. Balla JI. The use of critical cues and prior probability in decision-making. *Methods Inf Med* 1982; 21: 9-14.
3. Dolan JG, Bordley DR, Mushlin AI. An evaluation of clinicians' subjective prior probability estimates. *Med Decis Making* 1986; 6: 216-223.
4. Peeters A, Ting J, Nelson MR, McNeil JJ. Coronary heart disease risk prediction by general practitioners in Victoria [letter]. *Med J Aust* 2004; 180: 252.
5. Ayanian J, Berwick D. Do physicians have a bias toward action? *Med Decis Making* 1991; 11: 154-158.
6. Attia J, Balakrishnan RN, Sibbritt D, et al. Generating pre-test probabilities: a neglected area in clinical decision making. *Med J Aust* 2004; 180: 449-454.
7. Henrikson CA, Howell EE, Bush DE, et al. Chest pain relief by nitroglycerin does not predict active coronary artery disease. *Ann Intern Med* 2003; 139: 979-986.
8. Ingui BJ, Rogers MA. Searching for clinical prediction rules in MEDLINE. *J Am Med Inform Assoc* 2001; 8: 391-397.
9. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules: a review and suggested modifications of methodological standards. *JAMA* 1997; 277: 488-494.
10. McDonald CJ. Medical heuristics: the silent adjudicators of clinical practice. *Ann Intern Med* 1996; 124: 56-62. ❑