Users' Guides to the Medical Literature

VII. How to Use a Clinical Decision Analysis

B. What Are the Results and Will They Help Me in Caring for My Patients?

W. Scott Richardson, MD, Allan S. Detsky, MD, PhD, for the Evidence-Based Medicine Working Group

YOU RECALL from the first of our two articles concerning clinical decision analysis[1] that your patient is a middle-aged man with heart failure from an idiopathic dilated cardiomyopathy. You are trying to decide whether to recommend anticoagulation with warfarin to prevent systemic or pulmonary thromboembolism. Your literature search showed that no randomized clinical trials of warfarin for this use have been published. The search did discover a clinical decision analysis,[2] and in the first article, we showed you how to evaluate its validity. In this article, we will show you how to interpret the results and generalizability of a clinical decision analysis (Table).

From the Department of Medicine, University of Rochester (NY) School of Medicine and Dentistry (Dr Richardson), and the Departments of Health Administration and Medicine, University of Toronto (Ontario), and the Division of General Internal Medicine and Clinical Epidemiology, The Toronto (Ontario) Hospital (Dr Detsky).

A complete list of members (with affiliations) of the Evidence-Based Medicine Working Group appears in the first article of this series (*JAMA*. 1993;270:2093-2095). The following members contributed to this article: Gordon Guyatt, MD, MSc (chair); Deborah Cook, MD, MSc; Hertzel Gerstein, MD, MSc; Robert Hayward, MD, MPH; Anne Holbrook, MD, PharmD; Roman Jaeschke, MD, MSc; Elizabeth Juniper, MCSP, MSc; Mitchell Levine, MD, MSc; David Naylor, MD, DPhil; Andrew Oxman, MD, MSc; David Sackett, MD, MSc; Sean Tunis, MD, MSc; Stephen Walter, PhD; John Williams, Jr, MD, MHS; and Mark Wilson, MD, MPH.

Reprint requests to Room 2C12, McMaster University Health Sciences Centre, 1200 Main St W, Hamilton, Ontario, Canada L8N 3Z5 (Gordon Guyatt, MD, MSc).

As shown in the Figure, decision trees are displayed graphically, oriented from left to right, with the decision to be analyzed on the left, the compared strategies in the center, and the clinical outcomes on the right. The square box, termed a "decision node," represents the decision to be made, and the lines emanating from this decision node represent the clinical strategies being compared. Circles, or "chance nodes," represent chance events and outcome states are shown as triangles on the far right. Numbers beside the strategies (if they were present) would be "probabilities," or the likelihood of events, while the numbers by the outcome states would be "utilities," or the value of these events.[3,4]

In the Baseline Analysis, Does One Strategy Result in a Clinically Important Gain for Patients? If Not, Is the Result a Toss-up?

For a clinical decision analysis that compares two clinical strategies, there are three possible results: the first strategy is better than the second, the second strategy is better than the first, or both strategies are equally good (or equally bad), a result known as a "toss-up" or a "close call."[5] For instance, in an analysis of the management of solitary pulmonary nodules, the analysts found the choice of strategies to be a close call in terms of expected gains in life expectancy.[6] The larger the number of strategies compared in an analysis, the larger the number of possible results, but always with the same idea: any one strategy can "win" or two or more strategies could "tie." The terms "baseline" or "base case" refer to the set of numbers for probability that the analyst believes are closest to the actual state of affairs.

One chooses between strategies in a decision tree by comparing the overall benefits expected from pursuing each strategy, termed its "expected utility," and then selecting the strategy with the highest value of expected utility. Some controversy remains as to when exceptions to this rule are legitimate or desirable.[7] To calculate expected utility, one starts at the rightmost branches of the tree, multiplies the probability for each by its utility, and sums these products for each chance node. One repeats this calculation moving leftward, a process known as "folding back," until one has calculated the expected utility value for each strategy.

For example, consider the topmost chance node in the Figure, with its two branches. Imagine that the "no-embolism" and "embolism" branches have probabilities of 0.95 and 0.05 and utilities of 1.0 and .9, respectively. The expected utility for this chance node would be the sum of the

Users' Guides to the Medical Literature section editor: Drummond Rennie, MD, Deputy Editor (West), *JAMA*.

User's Guides for Clinical Decision Analysis

Are the results valid?
- Were all important strategies and outcomes included?
- Was an explicit and sensible process used to identify, select, and combine the evidence into probabilities?
- Were the utilities obtained in an explicit and sensible way from credible sources?
- Was the potential impact of any uncertainty in the evidence determined?

What are the results?
- In the baseline analysis, does one strategy result in a clinically important gain for patients?
  - If not, is the result a toss-up?
- How strong is the evidence used in the analysis?
- Could the uncertainty in the evidence change the result?

Will the results help me in caring for my patients?
- Do the probability estimates fit my patients' clinical features?
- Do the utilities reflect how my patients would value the outcomes of the decision?

product of each of the probabilities times the utilities, in this case $(0.95 \times 1.0) + (0.05 \times 0.9)$, which equals 0.995.

The decision analyst chooses the scale on which these expected utilities are measured to fit the clinical problem. For instance, in an analysis of strategies that could reduce death, the analyst might choose to measure utility as the number of lives saved or the average gain in remaining life expectancy, both measures of the quantity of life. Other utility scales can be used to report on the quality of life. Both quantity and quality can be combined into a single measure, such as quality-adjusted life years[8] or healthy-years equivalence.[9] For instance, suppose one strategy in a decision analysis yielded an average remaining life expectancy of 5 years, but that all five years were lived in a state of health rated by patients to have a utility value of 0.8. The quality-adjusted life expectancy would be 5×0.8 or 4 years.

Now that you understand where the results of the decision analysis come from, you must decide if any difference between strategies is clinically important. In making this judgment, consider that the differences presented will be average differences rather than differences that you can expect for every patient. Some patients will gain considerably more, while others will gain considerably less. This is no different than interpreting average differences between groups in randomized trials. You may not, however, be familiar with differences in life expectancy, the output of many decision analyses. Keep in mind that a gain in life expectancy does not occur just at the end of a person's life—it may occur at the beginning or be spread over the course of time.[10]

How large must a gain in remaining life expectancy be to be important? Probably smaller than you might think, although the answer to this question depends on judgments about several variables, and this controversial area has not yet been fully addressed by empirical research. In some recent studies, decision analysts have "translated" the results of clinical trials into life expectancy gains, for various widely accepted clinical interventions.[10,11] These studies suggest that a gain in life expectancy or quality-adjusted life expectancy of 2 or more months ought to be considered an important gain, while a gain of a few days would represent a toss-up.

In the anticoagulation for dilated cardiomyopathy example, the decision analysis finds warfarin to be the preferred strategy for all patients 35 to 75 years of age. The average gain in quality-adjusted life expectancy for 55-year-old patients (similar to your 51-year-old patient) is 115 days, or almost 3 months. From the above, you can see that this gain in life expectancy is probably important. Since the analysts explicitly considered both the reduction of emboli and the risk of bleeding, this 3-month gain in life expectancy represents the *net* clinical benefit you could expect from recommending anticoagulation to your patient.

How Strong Is the Evidence Used in the Analysis?

The probabilities used in clinical decision analyses are estimates, taken mostly from the published literature, and while they may represent the best available evidence, they are nonetheless subject to potential error. The best defense against such error is for the analysts to base probability estimates on studies of high methodological quality, after a thorough and unbiased search for all relevant studies. The analysts should explain how they judged the quality of these primary studies. One way to do this would be to judge study quality by applying criteria akin to those in the other articles in this series, whether for primary studies of therapy,[12,13] diagnosis,[14,15] harm,[16] prognosis,[17] or for integrative studies, such as overviews.[18]

As with other integrative studies, the overall strength of the result of a clinical decision analysis depends on the strength of inference possible from the primary studies. Ideally, every probability estimate at every node in the tree is supported by precise estimates from primary and integrative studies of high methodological quality, but such idealized analyses are rare. Good decision analyses can still be performed with some imprecise or ambiguous data, as long as most of the data are of good quality and the analysts explain any limitations and plan their sensitivity analyses accordingly. The fewer the probabilities that can be precisely estimated from high quality primary studies, ie, the weaker the evidence used in the analysis, the weaker the overall inference one can make from the results.

In the anticoagulation example, the authors describe vigorous efforts to obtain the correct values for probabilities from the published literature and from experts. They highlight the limited methodological quality of the primary literature and acknowledge the weakened inference. In particular, there are no randomized trials to tell you whether patients with cardiomyopathy will live longer or have fewer morbid events if given anticoagulants.

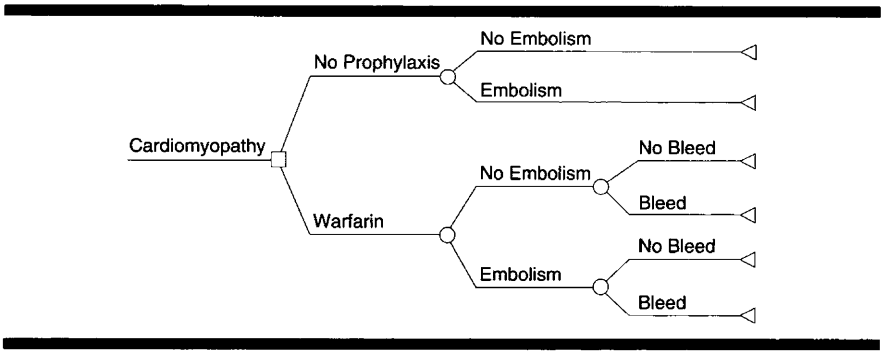
Could the Uncertainty in the Evidence Change the Result?

For any clinical variable such as the probability of bleeding, or the value that patients place on avoiding a stroke, the decision analyst can calculate the value, or "threshold," above which the results favor one strategy, and below which the results favor another strategy. For multiway sensitivity analyses, the analyst can show two-dimensional graphs of the variables, with the thresholds displayed as a line (two-way analyses) or a series of lines (three-way analyses) separating zones of strategy preference. While this may be daunting at first, these tables and graphs provide the most clinically useful information from a decision analysis.

If the result of the analysis (one strategy is preferred or a toss-up is found) would change by choosing different values for one of the variables, the result is said to be "sensitive" to that variable. On the other hand, if changing the variable throughout its plausible range of values doesn't change the result, the analysis result is said to be "robust" to the sensitivity analysis. As you might guess, the more robust the result is, the more confident you can be that the recommended strategy should indeed be preferred. If the result was a toss-up and that indifference proves robust to sensitivity analyses, you can be confident that the strategies are equivalent.

The analysts of the anticoagulation example found the preference for the warfarin strategy to be robust to the sensitivity analyses they completed, with two exceptions (we will return to one of these, the bleeding risk for patients taking warfarin).

For the other exception, the analysts assumed in the base case that patients' quality of life was not impaired by the inconvenience and anxiety associated with taking warfarin (ie, a utility value of 1.0 on a 0 to 1.0 scale). When testing



Structure of a decision tree. Square indicates decision node; circles, chance nodes; triangles, outcome nodes; and lines, strategy pathways. Numbers (when present) by lines indicate probabilities, and by triangles, utilities.

this assumption, by adjusting downward the utility rating for quality of life while taking warfarin, the analysts discovered that the choice of strategies would change substantially. For 55-year-old patients, the threshold utility value was 0.92. In other words, if patients rated their quality of remaining life while taking warfarin as 0.93 or greater, then anticoagulation would be preferred. For a utility rating of exactly 0.92, the two strategies would be equally preferred, while for utility ratings below 0.92, no anticoagulation would be preferred.

To put this result in perspective, recall that utility represents the value to the patient of remaining expected life, and that a rating of 0.92 is 8% less than normal. In other words, a utility threshold of 0.92 means that your patient feels he would be willing to sacrifice 8% of his remaining life to avoid taking warfarin. On a time scale, this means that a year taking warfarin would have to be worth only approximately 11 months of life not taking warfarin, in order for him to choose not to take it.

WILL THE RESULTS HELP ME IN CARING FOR MY PATIENTS?

Do the Probability Estimates Fit My Patients' Clinical Features?

This first issue of applicability concerns whether the clinical characteristics of patients for whom the analysis was intended are similar to your patients. For a decision analysis built for an individual patient, look for the description of that patient's condition; if the patient is well described, you should be readily able to judge how closely your patient resembles her or him. An article reporting a decision analysis built for a group of patients should have an analogous portion of the text, detailing the clinical characteristics of patients to whom the results of the analysis are to be applied. You should satisfy yourself that your patient would be included in this group.

You could be confident that the probabilities fit your patients if the estimates were taken from one or more rigorous clinical studies in which patient samples included patients similar to yours. If the authors don't describe the samples, you could track down the references and review the inclusion and exclusion criteria to see whether your patient would fit.

If the analysis was intended for patients different from yours, review the results of the sensitivity analyses. The clinical variables used for these analyses should be detailed enough for you to locate where your patient would fit, and thus what net benefit your patient might expect from the clinical strategies. If you still can't tell, ask yourself whether the clinical characteristics of the intended patients are so different from yours that you should discard the results. If not, you can proceed, with some caution, to use them.

In the anticoagulation example, most of the probabilities fit your dilated cardiomyopathy patient, including the rates of systemic and pulmonary emboli and the estimated mortality. The baseline average annual risk of major hemorrhage on warfarin was estimated to be 4.5%. If you worried that your patient's risk of bleeding while taking warfarin could be higher than average, you should examine the sensitivity analyses for this variable. These sensitivity analyses show that anticoagulation with warfarin remains the preferred strategy until the annual bleeding risk reaches 15%, more than triple the baseline estimate. Above this value, no anticoagulation became the preferred strategy.

When a clinical decision analysis shows that the preferred strategy is sensitive to a given variable, you will need to gauge where your patient fits on the scale of that variable. Thus, when deciding how to use the results of the anticoagulation decision analysis for your particular patient, you will need to estimate his annual risk of bleeding while undergoing warfarin therapy. While a full discussion of estimating the bleeding risk is beyond the scope of this article, we offer a few suggestions.

First, look in the text for the authors' description of their systematic review of the literature. Ideally, they will have found one or more original articles or systematic reviews of high methodological quality from which they obtained their baseline estimate, and from which you could obtain an individualized estimate for your patient. Alternatively, you could do your own search for this information, using the tactics introduced in the first article in this series.[19]

If you did so you would find a systematic review of this topic,[20] wherein the authors cite the average annual frequencies of fatal and major hemorrhage in patients taking warfarin as 0.6% and 3.0%, respectively. You might also find a study of warfarin use in atrial fibrillation,[21] wherein the incidence of major or fatal bleeding was 2.5%. If these numbers are close to the truth, then by using somewhat higher figures in the anticoagulation decision analysis, the analysts would have overestimated the risk of harm and might have obscured a net benefit. Despite this, the warfarin strategy still resulted in a clinically important expected gain in life expectancy, suggesting that the true net benefit might be somewhat larger than reported. Note also that these published estimates are substantially lower than the 15% threshold value for annual bleeding risk, above which the no-warfarin strategy would be preferred.

Your search would also turn up a retrospective analysis of thromboembolism rates in two randomized trials of other treatments (not anticoagulants) for heart failure.[22] During the approximately 2.5 years average follow-up, the trials showed thromboembolic events occurred in 4.7% and 5.2% of patients. After transformation to comparable event rates, these results may be a little over half of the values used in the anticoagulation decision analysis. By using somewhat higher estimates, the analysts could have overestimated the benefit of warfarin.

Do the Utilities Reflect How My Patients Would Value the Outcomes of the Decision?

Since the utility ratings for the value of outcomes has a strong influence on the choice of strategies, you must consider whether your patient's values are similar to those used in the decision analysis. In a decision analysis built for an individual patient, the utilities are usually measured directly from that patient, so while those values should be quite believable for that patient, they may not necessarily fit your patient. Alternatively, utilities measured from a

large group of patients or members of the general public would probably include a set of values similar to those of your patient, but the range of values might be so broad that you are left uncertain as to which values to use. If you encounter such difficulties, you should examine the one-way and multiway sensitivity analyses that use a wide range of utility estimates to see how your patient's values will affect the final decision.

If you were to ask your patient to rate the outcome states using the rating instrument in the article, you would know exactly what utility values to use. However, most clinicians won't have the time or inclination to do this. Fortunately, you can still make some judgment about this question by asking your patient about values in nonquantitative terms. For instance, one patient may be extremely averse to regular monitoring, while another may not mind. Disabling stroke might devastate one patient, whereas another might be more resilient.

As mentioned above in the anticoagulation example, the utility rating for life while taking warfarin had a substantial influence on the preference of strategies. The authors highlight the importance of this variable and urge that investigators examine patients' reactions to taking warfarin and undergoing monitoring, so that subsequent recommendations about anticoagulation can be better informed.

RESOLUTION OF THE SCENARIO

Without a randomized trial of anticoagulation in patients with dilated cardiomyopathy in sinus rhythm, your overall confidence in a decision to anticoagulate your patient will be limited. In the absence of trial data, experts have recommended that the decision to use warfarin in this setting be made on an individual basis.[23-25] How are you to individualize the treatment decision for your middle-aged man with dilated cardiomyopathy? The anticoagulation decision analysis suggests that if he has a low or moderate bleeding risk and a ready acceptance of anticoagulation monitoring, he is likely to be better off taking warfarin. Thus, the decision analysis identifies the few clinical variables on which the decision depends, and estimates the size and likelihood of net clinical benefit you could expect from the alternative courses of action. While the better therapy may still be unproved, you should now be much more informed about the choice and better prepared to decide with the patient what is to be done.

Dr Detsky is supported in part by a National Health Research Scholar Award from Health and Welfare Canada.

References

1. Richardson WS, Detsky AS, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, VII: how to use an article on clinical decision analysis, A: are the results of the study valid? *JAMA.* 1995;273:1292-1295.
2. Tsevat J, Eckman MH, McNutt RA, Pauker SG. Warfarin for dilated cardiomyopathy: a bloody tough pill to swallow? *Med Decis Making.* 1989;9:162-169.
3. Weinstein MC, Fineberg HV. *Clinical Decision Analysis.* Philadelphia, Pa: WB Saunders; 1980.
4. Sox HC, Blatt MA, Higgins, MC, Marton KI. *Medical Decision Making.* Boston, Mass: Butterworth-Heinemann; 1988.
5. Kassirer JP, Pauker SG. The toss-up. *N Engl J Med.* 1981;305:1467-1469.
6. Cummings SR, Lillington GA, Richard RJ. Managing solitary pulmonary nodules: the choice of strategy is a 'close call.' *Am Rev Respir Dis.* 1986;134:453-460.
7. Deber RB, Goel V. Using explicit decision rules to manage issues of justice, risk and ethics in decision analysis: when is it not rational to maximize expected utility? *Med Decis Making.* 1990;10:181-194.
8. Torrance GW, Feeny D. Utilities and quality-adjusted life years. *Int J Technol Assess Health Care.* 1989;5:559-579.
9. Mehrez A, Gafni A. Quality-adjusted life years, utility theory and healthy-years equivalence. *Med Decis Making.* 1989;9:142-149.
10. Naimark DM, Naglie G, Detsky AS. The meaning of life expectancy: what is a clinically significant gain? *J Gen Intern Med.* 1994;9:702-707.
11. Tsevat J, Weinstein MC, Williams LW, et al. Expected gains in life expectancy for various coronary heart disease risk factor modifications. *Circulation.* 1991;83:1194-1201.
12. Guyatt GH, Sackett DL, Cook DJ, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, II: how to use an article about therapy or prevention, A: are the results of the study valid? *JAMA.* 1993;270:2598-2601.
13. Guyatt GH, Sackett DL, Cook DJ, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, II: how to use an article about therapy or prevention, B: what were the results and will they help me in caring for my patients? *JAMA.* 1994;271:59-63.
14. Jaeschke R, Guyatt GH, Sackett DL, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, III: how to use an article about a diagnostic test, A: are the results of the study valid? *JAMA.* 1994;271:389-391.
15. Jaeschke R, Guyatt GH, Sackett DL, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, III: how to use an article about a diagnostic test, B: what are the results and will they help me in caring for my patients? *JAMA.* 1994;271:703-707.
16. Levine MS, Walter SS, Lee HN, Haines T, Holbrook A, Moyer V, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, IV: how to use an article about harm. *JAMA.* 1994;271:1615-1619.
17. Laupacis A, Wells G, Richardson WS, Tugwell P, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, V: how to use an article about prognosis. *JAMA.* 1994;272:234-237.
18. Oxman AD, Cook DJ, Guyatt GH, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, VI: how to use an overview. *JAMA.* 1994;272:1367-1371.
19. Oxman AD, Sackett DL, Guyatt GH, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, I: how to get started. *JAMA.* 1993;270:2093-2095.
20. Landefeld CS, Beyth RJ. Anticoagulant-related bleeding: clinical epidemiology, prediction and prevention. *Am J Med.* 1993;95:315-328.
21. Connolly SJ, Laupacis A, Gent M, Roberts RS, Cairns JA, Joyner C, for the CAFA Study Coinvestigators. Canadian atrial fibrillation anticoagulation (CAFA) study. *J Am Coll Cardiol.* 1991;18:349-355.
22. Dunkman WB, Johnson GR, Carson PE, et al. Incidence of thromboembolic events in congestive heart failure. *Circulation.* 1993;87[suppl VI]:VI94-VI101.
23. Dec GW, Fuster V. Idiopathic dilated cardiomyopathy. *N Engl J Med.* 1994;331:1564-1575.
24. Baker DW, Wright RF. Management of heart failure, IV: anticoagulation for patients with heart failure due to left ventricular systolic dysfunction. *JAMA.* 1994;272:1614-1618.
25. Kubo SH, Cohn JN. Approach to treatment of the patient with heart failure in 1994. *Adv Intern Med.* 1994;39:485-515.