Users' guides to the medical literature
Richardson, W Scott; Detsky, Allan S
JAMA; Apr 26, 1995; 273, 16; AMA Titles
pg. 1292

Users' Guides to the Medical Literature

VII. How to Use a Clinical Decision Analysis

A. Are the Results of the Study Valid?

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

CLINICAL SCENARIO

You are the attending physician on an inpatient service where a 51-year-old man is admitted with congestive heart failure of recent onset. You find he has a dilated cardiomyopathy, the cause of which remains unknown after a thorough evaluation. He is in sinus rhythm. The team's resident asks you whether the patient should be anticoagulated with warfarin, enough to keep his international normalized ratio from 2.0 to 3.0, in order to prevent systemic emboli, even though his echocardiogram does not show left ventricular thrombus. You are not sure about the evidence concerning this issue, so you admit your shared knowledge gap and resolve to search together for the relevant information.

THE SEARCH

In the hospital's library, the two of you search the MEDLINE system using several search terms, such as "cardiomyopathy, dilated," "cardiomyopathy, congestive," and "heart failure, congestive" crossed with "warfarin," "anticoagulation," and "thromboembolism." Despite several attempts, you retrieve no randomized trials of warfarin used for this purpose. Even after enlisting the help of the librarian, you are unable to locate any clinical trials about this question. You do come across an editorial calling for a clinical trial of your question.[1] You also retrieve two review articles, one that recommends anticoagulation for such patients,[2] and the other that recommends no anticoagulation.[3] The latter review cites a decision analysis on this issue,[4] which you retrieve, hoping to find further guidance for your decision.

INTRODUCTION

Decision making involves choosing an action after weighing the risks and benefits of the alternatives. While all clinical decisions are made under conditions of uncertainty, the degree of uncertainty decreases when the medical literature includes directly relevant, valid evidence. When the published evidence is scant, or less valid, uncertainty increases.

Decision analysis is the application of explicit, quantitative methods to analyze decisions under conditions of uncertainty. Decision analysis allows clinicians to compare the expected consequences of pursuing different strategies. The process of decision analysis makes fully explicit all of the elements of the decision, so that they are open for debate and modification. While a decision analysis will not solve your clinical problems, it can help you explore the decision.[5-7]

We will use the term "clinical decision analyses" to include studies that analyze decisions faced by clinicians in the course of patient care, such as deciding whether to screen for a condition, choosing a testing strategy, or selecting a treatment. While such analyses can be undertaken to inform a decision for an individual patient ("Should I recommend warfarin to this 51-year-old man with idiopathic dilated cardiomyopathy?"), they are more widely undertaken to help inform a decision about clinical policy[8] ("Should I routinely recommend warfarin to patients in my practice with dilated cardiomyopathy?"). The study retrieved by the search for our scenario is an example of this latter type, while an example of the former is the analysis by Wong et al[9] of whether to recommend cardiac surgery for an elderly woman with aortic stenosis.

Decision analysis can also be applied to more global questions of health care policy, analyzed from the perspective of society or a national health authority. Examples include analyses of whether or not to screen for prostate cancer[10] and comparing different policies for cholesterol screening and treatment.[11] While decision analyses in health services research share many attributes with clinical analyses,[12] they are sufficiently different that they are beyond the scope of these articles.

In helping you understand decision analysis, we will review some of the "anatomy and physiology" of decision models. This is not meant to be an ar-

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, MSc; Roman Jaeschke, MD, MSc; Elizabeth Juniper, MCSP, MSc; Mitchell Levine, MD, MSc; David Naylor, MD, DPhil; Andrew Oxman, MD, MSc, FACPM; 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).

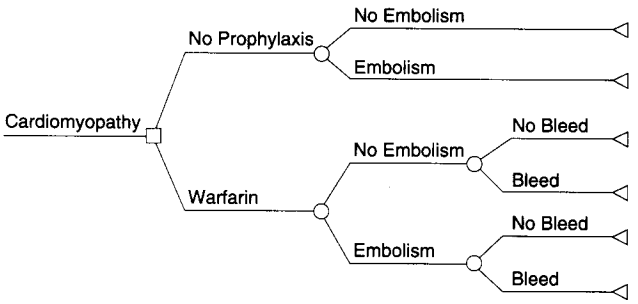
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Users' 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?



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.

ticle on how to perform decision analysis; if you wish to read about that, you should look elsewhere.[13,14]

FRAMEWORK FOR THE USERS' GUIDES

We will approach articles on clinical decision analysis using the same framework introduced in earlier articles in this series, as follows:

Are the Results Valid?

This question addresses whether the strategy recommended by the analysis is truly likely to be the better one for patients. Just as with other types of studies, the validity of a decision analysis is largely determined by the strength of the methods used.

What Are the Results?

The users' guides under this second question consider the size of the expected net benefit from the recommended strategy and our confidence in this estimate of net benefit.

Will the Results Help Me in Caring for My Patients?

If the decision analysis yields valid and important results, you should examine whether these results can be generalized to the patients in your practice.

The Table summarizes the specific guides you should use when addressing these three questions. We will explore the guides by applying them to the study we found in our search. This article will deal with the validity guides, while the next in the series will address the results and applicability.

ARE THE RESULTS VALID?

Were All Important Strategies and Outcomes Included?

At issue here is how well the structure of the model fits the clinical decision you face. Most clinical decision analyses are built as decision trees, and the articles will usually include one or more diagrams showing the structure of the decision tree used for the analysis. Reviewing these diagrams will help you understand the model. You must then judge whether the model fits the clinical problem well enough to be valid.

The Figure shows a diagram of a much simplified version of the decision tree for the anticoagulation problem. The clinician has two options for patients with cardiomyopathy, either to offer no prophylaxis or to prescribe warfarin. Either way, patients may or may not develop embolic events. Prophylaxis lowers the chance of embolism but can cause bleeding in some patients. As seen 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 decision is diagrammed by a square, termed a "decision node." The lines emanating from the decision node represent the clinical strategies being compared. Chance events are diagrammed with circles, called "chance nodes," and outcome states are shown as triangles or as rectangles.

To explore more fully how the model's structure affects its validity, we will highlight two aspects here.

Were All of the Realistic Clinical Strategies Compared?—In a decision analysis, a strategy is defined as a sequence of actions and decisions that are contingent on each other. For instance, the strategy of anticoagulant therapy for a patient includes not only the prescription and the monitoring, but also the adjustment of the warfarin dose for changes in prothrombin time. The authors should specify which decision strategies are being compared (at least two, otherwise there's no decision). Further, the clinical strategies included should be described in enough detail to recognize them as separate and realistic choices. You should satisfy yourself that the clinical strategies you consider important are included in the analysis.

For example, in a decision analysis of the management of suspected herpes encephalitis, the authors included the three strategies available to clinicians then: brain biopsy, empirical vidarabine, or neither.[15] At that time, this model represented the clinical decision well. Since then, however, acyclovir has become available and has been widely used for this disorder. Because the original model did not include an acyclovir strategy, it would no longer accurately portray the decision.

In the anticoagulation example, the analysts studied two clinical strategies, warfarin and no warfarin. This fits quite well the clinical decision you face in the scenario. Note that the decision model does not include a third strategy of using aspirin instead of warfarin. If, when considering the treatment options for this patient, you would seriously consider the use of aspirin instead of warfarin, then you would judge this model as incomplete.

Were All Clinically Relevant Outcomes Considered?—To be useful to clinicians and patients, the decision model should include the outcomes of the disease that matter to patients. Generally speaking, these include not only the quantity of life but also its quality, in measures of disease and disability. Obviously, the specific disorder in question determines which outcomes are clinically relevant. For an analysis of an acute, life-threatening condition, life expectancy might be appropriate as the main outcome measure. But in an analysis of diagnostic strategies for a nonfatal disorder, more relevant outcomes would be discomfort from testing or days of disability avoided. By examining the outcomes used in the analysis, you can discover the viewpoint from which the analyst built the decision model. Clinical decision analyses should be built from the perspective of the patient, that is, should include all the clinical benefits

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and risks of importance to patients (they can include other considerations as well).

Also, by comparing the outcomes between strategies, you can discover the trade-offs built into the model. Most clinical dilemmas are dilemmas because they include trade-offs between competing benefits and competing risks. For instance, when deciding how best to manage small abdominal aortic aneurysms, one must weigh reducing the risk of aneurysm rupture against the chance of unnecessary surgery in patients who would have died from other causes before rupture.[16] For a decision analysis to be worth doing, ie, for the clinical decision to be difficult enough, the choice of strategies should be balanced on one or more of such trade-offs. You should satisfy yourself that these important trade-offs are represented well in the model's structure.

For the anticoagulation example, the authors' decision model includes all of the clinical events of interest to patients (stroke, other emboli, hemorrhage, and the like). The outcomes are measured as "quality-adjusted life expectancy," a scale that combines information about both the quantity and the quality of life. This metric fits your clinical decision well, for you can expect that warfarin might affect both the quantity and quality of life. By reviewing the tree diagram, you can see that the authors have included the principal trade-off in the decision: the warfarin strategy offers the benefits of preventing systemic arterial embolism causing stroke and preventing pulmonary embolism, while it could cause the harm of bleeding.

Was an Explicit and Sensible Process Used to Identify, Select, and Combine the Evidence Into Probabilities?

To assemble the large amount of information necessary for a decision analysis, the analyst searches the published literature and interviews experts and patients. Just as with other integrative studies like overviews,[17] authors of clinical decision analyses should search and select the literature in an explicit and unbiased way, and then appraise the validity, effect size, and homogeneity of the studies in a reproducible fashion. Ideally, they would judge study quality by applying criteria akin to those in the other articles in this series, whether for primary studies of therapy,[18,19] diagnosis,[20,21] harm,[22] prognosis,[23] or for other integrative studies, such as overviews.[17] In other words, the authors should perform as comprehensive a literature review as is required for a meta-analysis.

Once gathered, the information must be transformed into quantitative estimates of the likelihood of events, or probabilities. The scale for probability estimates ranges from 0 (impossible) to 1.0 (absolutely certain). Probabilities must be assigned to each branch emanating from a chance node, and for each chance node, the sum of probabilities must add to 1.0.

For example, looking at the Figure, note that the no-anticoagulation strategy (the upper branch coming from the decision node) has one chance node, at which two possible events could occur, either an embolism or no embolism (labeled "no embolism"). To assign a probability to these two branches from the chance node, the analyst tracks down all relevant evidence about the rates of systemic emboli in patients with cardiomyopathy. If the best estimate of the rate were found to be 5%, then the analyst would assign 0.05 to the embolism branch and 0.95 to the no-embolism branch.

Usually, rates from clinical studies can be directly translated into probabilities, as in this example. In other instances, the data must be transformed first, such as when analysts must adjust 5-year survival data to fit an analysis concerned with only the first 3 years. Analysts should report which data were used and how the data were transformed.

In the anticoagulation example, the authors describe vigorous efforts to obtain the correct values for probabilities from the published literature and from experts, although they don't provide the search terms they used. The authors do highlight the limited data available and the data's methodological limits. Also, they tabulate the evidence they use and mention the transformations needed for the model.

Were the Utilities Obtained in an Explicit and Sensible Way From Credible Sources?

Utilities represent quantitative measurements of the value to the decision maker of the various outcomes of the decision. Several methods are available to measure these values directly,[5,7,24,25] and which method is best remains controversial. Different methods use different scales; a commonly used utility scale ranges from 0 (worst outcome, usually death) to 1.0 (excellent health). Whatever the measurement method used, the authors should report the source of the ratings. In a decision analysis built for an individual patient, the most (and probably only) credible ratings are those measured directly from that patient. For analyses built to inform clinical policy, credible ratings could come from three sources: (1) direct measurements from a large group of patients with the disorder in question and to whom results of the decision analysis could be applied; (2) from published studies of quality-of-life ratings by such patients, as was done in a recent analysis of strategies for chronic atrial fibrillation[26]; or (3) from an equally large group of people representing the general public. Whoever provides the rating must understand the outcomes they are asked to rate; the more the raters know about the condition, the more credible are their utility ratings.

The authors of the anticoagulation example obtained values from several internists familiar with the clinical disorder and with the treatments. While physician raters were undoubtedly familiar with the outcomes of systemic emboli and major hemorrhage, only a small number of physicians made ratings, and their values may not represent those of either patients or the general public.

Was the Potential Impact of Any Uncertainty in the Evidence Determined?

Much of the uncertainty in clinical decision making arises from the lack of valid evidence in the literature. This lack of data hampers both clinical decision making and formal decision analysis. Even when it is present, published evidence is often imprecise, with wide confidence intervals around estimates for important variables. For instance, in a decision analysis concerning the management of polymyalgia rheumatica, the analysts searched the literature for the test sensitivity of temporal artery biopsy for giant cell arteritis.[27] The reported test sensitivity ranged from about 60% to 100%. In the decision analysis, these analysts set the baseline value equal to 83%, but repeated the analysis for values between 60% and 100%.

Decision analysts use this systematic exploration of the uncertainty in the data, known as "sensitivity analysis," to see what effect varying estimates for risks, benefits, and values have on the expected clinical outcomes, and therefore on the choice of clinical strategies. Sensitivity analysis asks the question: is the conclusion generated by the decision analysis affected by the uncertainties in our estimates of the likelihood or value of the outcomes? Estimates can be varied one at a time, termed "one-way" sensitivity analyses, or two or three at a time, known as "multi-way" sensitivity analyses. You should look for a table listing which variables were included in the sensitivity analyses, what range of values were used for each variable, and which variables, if any, altered the choice of strategies. Sat-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

isfy yourself that all of the clinically important variables were examined.

Generally, all of the probability estimates should be tested using sensitivity analyses. The range over which they should be tested will depend on the source of the data. If the estimates come from large, high-quality randomized trials with narrow confidence limits, the range of estimates tested can be narrow. The less valid the methods, or the less precise the estimates, the wider the range that must be included in the sensitivity analyses.

Utility values should also be tested with sensitivity analyses, with the range of values again determined by the source of the data. If large numbers of patients or knowledgeable and representative members of the general public gave very similar ratings to the outcome states, a narrow range of utility values can be used in the sensitivity analyses. If the ratings came from a small group of raters, or if individuals varied widely in their values, then investigators should use a wider range of utility values in the sensitivity analyses.

In the anticoagulation example, the authors responded to the poor quality of their evidence by varying all of the important variables over wide ranges. They report the results from several, although not all, of these sensitivity analyses, including the effect of higher bleeding risk while taking warfarin.

In the next article on clinical decision analysis, we will show you how to determine what the results are and how to use them in your practice.

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

References

1. Falk RH. A plea for a clinical trial of anticoagulation in dilated cardiomyopathy. *Am J Cardiol.* 1990;65:914-915.
2. Dec GW, Fuster V. Idiopathic dilated cardiomyopathy. *N Engl J Med.* 1994;331:1564-1575.
3. 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.
4. 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.
5. Keeney RL. Decision analysis: an overview. *Operations Res.* 1982;30:803-838.
6. Eckman MH, Levine HJ, Pauker SG. Decision analytic and cost-effectiveness issues concerning anticoagulant prophylaxis in heart disease. *Chest.* 1992;102(suppl 4):538S-549S.
7. Kassirer JP, Moskowitz AJ, Lau J, Pauker SG. Decision analysis: a progress report. *Ann Intern Med.* 1987;106:275-291.
8. Eddy DM. Designing a practice policy: standards, guidelines, and options. *JAMA.* 1990;263:3077, 3081, 3084.
9. Wong JB, Salem DN, Pauker SG. You're never too old. *N Engl J Med.* 1993;328:971-975.
10. Krahn MD, Mahoney JE, Eckman MH, Trachtenberg J, Pauker SG, Detsky AS. Screening for prostate cancer: a decision analytic view. *JAMA.* 1994;272:773-780.
11. Krahn MD, Naylor CD, Basinski AS, et al. Comparison of an aggressive (U.S.) and a less aggressive (Canadian) policy for cholesterol screening and treatment. *Ann Intern Med.* 1991;115:248-255.
12. Goel V. Decision analysis: applications and limitations. *Can Med Assoc J.* 1992;147:413-417.
13. Weinstein MC, Fineberg HV, et al. *Clinical Decision Analysis.* Philadelphia, Pa: WB Saunders; 1980.
14. Sox HC, Blatt MA, Higgins MC, Marton KI. *Medical Decision Making.* Boston, Mass: Butterworth-Heinemann; 1988.
15. Barza M, Pauker SG. The decision to biopsy, treat, or wait in suspected herpes encephalitis. *Ann Intern Med.* 1980;92:641-649.
16. Katz DA, Littenberg B, Cronenwett JL. Management of small abdominal aortic aneurysms: early surgery vs watchful waiting. *JAMA.* 1992;268:2678-2686.
17. 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.
18. 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.
19. 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.
20. 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.
21. 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.
22. 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.
23. 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.
24. Llewellyn-Thomas H, Sutherland HJ, Tibshirani R, et al. The measurement of patient values in medicine. *Med Decis Making.* 1982;2:449-462.
25. Dolan JG, Isselhardt BJ, Cappuccio JD. The analytic hierarchy process in medical decision making: a tutorial. *Med Decis Making.* 1989;9:40-50.
26. Disch DL, Greenberg ML, Holzberger PT, et al. Managing chronic atrial fibrillation: a Markov decision analysis comparing warfarin, quinidine, and low-dose amiodarone. *Ann Intern Med.* 1994;120: 449-457.
27. Buchbinder R, Detsky AS. Management of suspected giant cell arteritis: a decision analysis. *J Rheumatol.* 1992;19:1220-1228.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.