EVIDENCE-BASED CLINICAL PRACTICE

Evaluating Evidence From A Decision Analysis

Scott D. Ramsey, MD, PhD

A 28-year-old man comes to your office with symptoms, or morning stiffness. Each attack lasts about 2 to 3 weeks, then slowly subsides. Last summer, the patient went on a week-long backpacking trip in the mountains. He does not recall seeing a tick on his person or his clothing during the trip, nor did he notice a rash during his outing or after he returned. You find nothing remarkable during his physical examination. You consider Lyme disease in your differential diagnosis, remembering that one of your colleagues treated a patient empirically for Lyme disease early last fall when that patient had a rash that resembled erythema migrans. You are considering testing the patient for Lyme disease but are uncertain about the best approach to take, so a brief review of the literature is in order.

Searching the literature, you come across a decision analysis of test-treatment strategies for patients with suspected Lyme disease. The article describes three common clinical scenarios where Lyme disease is considered in the differential diagnosis: (1) myalgic symptoms, (2) rash resembling erythema migrans, and (3) recurrent oligoarticular inflammatory arthritis. For inflammatory arthritis symptoms (your patient), the article recommends no testing or empirical treatment for Lyme disease. Because this article is a decision analysis and not a report from a high-quality clinical trial, you question whether the findings are valid and should guide your clinical approach to the patient.

Evaluation of Decision Analysis

Practitioners of evidence-based medicine seek evidence from high-quality studies (such as randomized, controlled trials) to inform the decisions they make in day-to-day practice. Still, more often than not, direct evidence is unavailable to guide clinical decision making. Decisions must then be made with only partial or incomplete evidence. In these cases, one would still like to make a decision based on the best possible estimate of the impact of the therapy on outcomes rather than one based on a best guess or the vague concept of a standard of care.

Decision analysis is the systematic, quantitative approach to address clinical decision making under uncertainty. In theory, it is an extremely attractive way to address these situations. It explicitly integrates the best external evidence, clinical expertise, and individual patient choice to help inform decisions. It can improve decision making when the database is incomplete. Decision analysis has been used to guide decisions for individual patients and policy development for managing groups of patients, and often it is the first step in a cost-effectiveness analysis. As a result, many advocate more widespread use of decision analysis in all aspects of medicine. In practice, however, there are many barriers to using decision analysis, both at the bedside and in guiding clinical policy.

This article is directed toward readers who have considered the merits of decision analyses but who are not particularly familiar with the techniques of this discipline. First, it reviews the strengths and limitations of clinical decision analysis. Next, it points out several areas where decision analysis may be of most use to practicing clinicians. Finally, it describes the key elements that readers should look for to help them determine whether the decision analysis has been performed properly. Excellent reference articles and texts are available for those who would like to read further about this technique.

Strengths of Decision Analysis

Decision analysis can be a powerful tool to address questions that have clinical relevance but have not yet been subjected to high-quality clinical trials. In some cases, such as when controlled trials are not...
feasible (eg, trials of pacemakers for life-threatening arrhythmias), decision analysis might be the only way to address the problem systematically. In other cases, when the risks and benefits of the intervention are uncertain (perhaps because preliminary studies have yet to be published), decision analysis might be the only way to address the problem systematically. Third, even when definitive studies have been done on one population, they do not directly address the relative benefits of the intervention for a related but somewhat dissimilar population of interest.

For example, lowering cholesterol levels with 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase inhibitors has clearly been shown to be effective in reducing morbidity and mortality for patients with known coronary artery disease and elevated low-density lipoprotein (LDL) cholesterol. Even so, there is less strong evidence of the degree of benefit in those whose only substantial risk for coronary artery disease is an elevated LDL cholesterol level. Decision analysis can weigh the risk, benefits, and costs of interventions of cholesterol lowering in populations where definitive studies have not been done. Finally, decision analysis can provide timely information about a clinical problem that otherwise would take years to address through the mechanism of a clinical trial. In the example above, a test of cholesterol reducing-agents in those with no known coronary disease could take 10 to 20 years before showing significant morbidity or mortality differences in the treatment and control groups.

Although decision analysis cannot definitively resolve clinical controversies where data from clinical trials are unavailable, it can convey several important pieces of information regarding an intervention for a particular clinical problem. First, it can determine whether a beneficial effect is likely. Second, it can determine the likely degree of benefit. Third, it can detect holes in the evidence chain. Holes refer to points in the clinical management pathway where the evidence is weak or lacking. When a decision analysis detects such holes in the evidence chain, it can be illuminating and a force for funding high-quality clinical trials. For example, although, DEXA nuclear medicine scans are sensitive and specific for detecting osteoporosis, and there is evidence that treatment of osteoporosis in women reduces the incidence of hip and vertebral fractures, there is no direct evidence that screening asymptomatic women with the DEXA technology improves outcomes compared with usual practice. A cost-effectiveness study that contained a decision analysis found this weak link in the evidence chain and called for further study of the issue.

**Limitations of Decision Analysis**

Despite the real and potential advantages of decision analysis, even ardent advocates of the discipline will admit that its use for clinical decision making and policy analysis has been limited for two important reasons. With regard to clinical decision making, decision analyses are much more time consuming to create than those who are unfamiliar with the discipline usually are aware. Well-built decision analyses of complex problems can take months to years to complete. Even simple problems can take 2 to 3 days of work, even for those who are trained. As a result, ad hoc decision analyses have not been adopted for day-to-day clinical practice and probably will not be in the future. Rather, the primary role of decision analysis will probably be to inform clinical practice policy, such as whether to test dyspeptic patients for *Helicobacter pylori*.

Another limitation of decision analyses is that they usually require integrating pieces of information from different studies, none of which directly addresses the problem at hand. The information can vary widely, not only in quality, but also in applicability to the clinical problem. O'Brien has referred to this issue as the Frankenstein monster problem, because the analyst patches together the model with disparate information, hoping it will nevertheless behave in a predictable way. In addition, when holes in the evidence chain are found, analysts often must rely on educated guesses or expert opinion to complete one or more sections of the analysis. This surrogate information, however, is exactly the type that evidence-based medicine seeks to avoid! As a result, decision analysis (like medicine) will always be part art and part science. Nevertheless, adhering to certain principles will vastly improve the quality of a decision analysis.

Readers of these studies should assure themselves that these issues are in order before accepting the results of these studies.
Table 1. Three Patient Scenarios for Lyme Disease.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Presentation</th>
<th>Clinical Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Myalgic symptoms</td>
<td>Fatigue</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stiffness</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diffuse muscular aches</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tenderness</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Slowly expanding rash</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Malaise</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fatigue</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Intermittent fever</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Headache</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mild stiff neck</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Arthralgia or myalgia</td>
</tr>
<tr>
<td>B</td>
<td>Rash resembling erythema migrans</td>
<td>Recurrent attacks of marked painful swelling</td>
</tr>
<tr>
<td></td>
<td></td>
<td>One or more large joints affected</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Episodes last 2 wk at a time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Episodes occur every 3 mo</td>
</tr>
<tr>
<td>C</td>
<td>Recurrent oligoarticular inflammatory arthritis</td>
<td>Long episodes of complete remission</td>
</tr>
</tbody>
</table>

Adapted from Nichol et al.¹

Critical Appraisal of a Decision Analysis

In providing a guide for evaluating a decision analysis, it is assumed that the reader is searching for evidence that can help guide the management of a particular clinical problem for a given type of patient. For example, in the clinical problem at the start of the chapter, the physician who sees a patient with recurrent oligoarticular inflammatory arthritis needs to know whether obtaining serologic tests for Lyme disease is warranted. A search of the literature found an article entitled, "Test-Treatment Strategies for Patients Suspected of Having Lyme Disease: A Cost-Effectiveness Analysis."¹

The article contains a decision-analysis-type evaluation of the problem. The key questions this physician must ask of the study are as follows: (1) Does the study apply to my patient or patient group? (2) Are the methods used to address the problem appropriate? (3) Do the results help me improve the care of my patient or patient group compared with what I would have done before I read the article? These questions will be addressed in the sections below.

Does the Study Apply to My Patient or Patient Group?

This fundamental question should be addressed before reading further into the article. There are two elements to this question: the patient and the clinical problem. Because the authors of a decision analysis essentially fabricate the patient population and clinical situation, it is critical that the patient and clinical scenario be defined clearly and precisely. The patient group that is the subject of the decision model should be well-described and as specific as possible. Age range, sex, and medical history should be described whenever changing these details would alter the testing or treatment pathway. The clinical history and examination findings (if necessary) should be detailed in a way that is easily recognizable to the reader or clinician. In the article on Lyme disease, three patient scenarios were detailed in tabular format (Table 1). If however, one of the patient scenarios outlined in the article on Lyme disease was "presents with rash," this history is so vague and nonspecific that it would be difficult to know whether it applied to the situation that the reader faces.

Are the Methods Used to Address the Problem Appropriate?

After describing the patient, the clinical history, and examination findings, the details of the analysis must be described clearly and systematically. First, the decision pathway must be clearly explained and justified. This pathway is usually represented graphically in the form of a decision tree. The decision pathway usually begins with two or more options that the clinician can choose. On the decision tree, these options are represented by a small
Compliant

Not compliant

Therapy Subtree

Major side effect

Minor side effect

No side effect

Cardiac complication

Arthritic complication

Neurologic complication

No complication

Outcome Subtree

Lyme disease

No Lyme disease

Figure 1. Decision subtrees for Lyme disease.

square, called a decision node. In the Lyme disease example, there are four initial options: (1) treat all, (2) test using enzyme-linked immunosorbent assay (ELISA), (3) two-step testing, and (4) no testing or treatment (Figure 1). The options should be mutually exclusive, include all choices that are typical for the situation, and if applicable, include the new (atypical) option of interest.

A chain of events on the decision tree follows the decision node. The events begin with two or more outcomes, each of which occurs with a given probability. A small circle called a chance node represents the range of outcomes. The outcomes following chance node should include all that are possible for the patient; in other words, the probabilities for all outcomes included after the chance node should sum to 1. In the Lyme disease example, choosing the ELISA testing option brings the decision maker to a chance node with two outcomes: positive or negative (note that if an equivocal outcome was possible for this test, then this chance node would be incomplete). The outcomes can be intermediate, in which case other decision nodes or chance nodes follow, or final, in which they are endpoints in the clinical pathway. A final outcome is usually referred to as a terminal node. In the example, terminal nodes include no Lyme disease and Lyme disease with one of four sequelae: cardiac, arthritic, neurologic, or no complications. The total group of terminal nodes should be mutually exclusive and represent all major endpoints related to the condition of interest. Note that the outcomes for the patient who has Lyme disease are not exclusive. This result is problematic because values must be attached to each outcome (one of the next issues for evaluating the model). If the outcomes are not comprehensive and exclusive, then it is difficult for the decision-maker to judge whether including the missing outcomes would change the preferred choice for the initial decision node.

The reader who is satisfied that the decision tree represents all reasonable choices and outcomes for the clinical problem should next turn attention to the values that have been assigned to the probabilities and outcomes in the model. These values inform the model. As with other areas of evidence-
based medicine, the analysts must search the literature and select data from the highest quality studies. The process of gathering the information should be described explicitly, and citations should be provided for all values used in the model. Where the choice of a best value is uncertain (perhaps because of conflicting results from separate studies), the analyst should take care to justify the choice of the value used in the decision analysis. In some cases, the analyst will choose the result from one study instead of another. In other cases, results from two or more studies will be pooled. Ideally, this pooling would use an accepted technique such as meta-analysis. 19 , 20 In the Lyme disease study, the authors pooled two studies of the sensitivity and specificity of the diagnostic tests using a random effects model that is commonly used in meta-analyses. 21 In other cases, high-quality evidence for a particular input might not be available, and the analyst is forced to rely on expert opinion or personal judgment to inform the model. If so, this choice should be stated explicitly, as with data from other studies.

Frequently values assigned to the outcome nodes are scalar rankings of patient preferences, known as utilities, ranging from 0 (death) to 1 (optimal health). Utilities are commonly combined with life expectancy to derive quality-adjusted life year estimates. The utilities can be obtained from the literature as with the other data or from surveys conducted by the analysts. Common subjects for the surveys are patients, their relatives, laypersons, or experts familiar with the condition of interest. A survey of experts using a utility elicitation technique known as the time tradeoff was used for the Lyme disease study. In some cases, particularly unfavorable outcomes, such as time spent in the intensive care unit, are arbitrarily assigned a utility weight of zero. In any case, the reader should be satisfied that the valuation of the outcomes is credible. Credibility is, of course, a judgment call. The author judges credibility based on: (1) the size of the sample from which the utility values were elicited (smaller samples are suspect because they are more likely to vary from the true average for the population); (2) whether the utility weights match the author’s clinical experience with similar patients; and (3) whether the estimates are conservative; that is, are chosen to bias the outcome of the study away from the new or experimental intervention of interest.

Costs are frequently included as a part of decision analyses. As with other sources of data, they must be credibly valued. In general, the costs should reflect the value of the resources used rather than what is charged for those resources. Resource valuation is particularly important for interventions that will potentially be applied to a great many persons (such as screening tests or immunizations). Inappropriate over- or under-pricing can adversely influence policy decisions regarding these technologies.

A final issue that the reader of a decision analysis must evaluate is how the analysts address uncertainty in the model. Uncertainty can be an issue for the structure of the decision tree itself or in the estimates for the probabilities and outcomes that are used to inform the tree. Decision analysts systematically explore uncertainty in the model with a technique called sensitivity analysis. The basic approach involves varying a particular parameter value (such as the cost of a test) between the best case (low cost) and a worst case (very high cost) and then running the calculations again to determine how the variation affects the ultimate outcome of the model. Inputs can be varied individually (one-way sensitivity analysis) or simultaneously (two-through n-way sensitivity analysis). In practice, it is very hard for a reader to interpret the sensitivity analysis when more than 2 inputs are varied simultaneously.

A newer technique for estimating uncertainty involves varying several parameters simultaneously, given a probability distribution for all inputs. Multiple iterations are performed using a technique such as Monte Carlo simulation until a confidence interval is created around the estimate. 22 The analyst can then state the degree of confidence in the results in probabilistic manner, such as, “there is 5% probability that the true cost-effectiveness of the new treatment is greater than $100,000.”

Ultimately sensitivity analysis is part art as well as science. Thus, evaluating the robustness of the sensitivity analysis involves judgment. The author follows a few rules to judge the sensitivity analysis of a decision model. These rules involve looking for (1) whether all clinically important variables were subjected to sensitivity analysis; (2) whether the best and worst cases met or exceeded the author’s own judgment of the range of possible outcomes for the parameter; (3) whether more than one form of sensitivity analysis (eg, one-way and probabilis-

Decision Analysis Evidence 399
tic) was used to evaluate the estimate for parameters that alter the outcome of the analysis when varied; (4) whether tables or figures display the range of results for parameters where varying the results alters the outcome of the analysis.

The Lyme disease model used only one-way sensitivity analysis of the parameters in the model. The clinical variables were varied as far as the range of values that were recorded in published studies. Because most of the cited studies had relatively few patients, the range is probably too narrow to be fully robust. The authors noted that extreme values for the sensitivities and specificities of the diagnostic tests influenced which strategy was the most economically attractive for certain scenarios but did not provide graphical or tabular data showing which scenarios change.

Do the Results Help Me Improve the Care of My Patient or Patient Group Compared With What I Would Have Done Before I Read the Article?

Clinicians read the literature to improve the care they deliver to their patients. A decision analysis should provide information that will allow readers to improve the way they diagnose and manage health problems. Thus, even if the methods are sound, the study will not be useful unless the results and the implications of the results are described in a way that provides useful, practical information.

There are three issues to consider when evaluating the clinical usefulness of a decision analysis: (1) whether the recommended strategy for addressing the clinical problem is clear and unequivocal, (2) the magnitude of the difference between the recommended strategy and the next best alternative, (3) whether the recommended strategy is feasible in the reader's clinical practice.

Compared with simpler guidelines, complex, multilayered clinical practice guidelines are difficult to adhere to in practice. Similarly, the reader of a decision analysis needs a clearly stated bottom line to act on it in clinical practice. The recommendation should include a description of the patient, the clinical problem, and the preferred management strategy in relation to the alternatives. A rationale for the recommended strategy should be given (eg, less expensive, fewer morbid endpoints, or greater life expectancy or quality-adjusted life expectancy). This rationale should be described in the space of one or two short paragraphs. The authors of the Lyme disease paper do a nice job of succinctly stating their findings. For example, they state rationales for two of the three clinical scenarios outlined in Table 1:

For a patient with myalgic symptoms or oligo-articular arthritis, two-step testing was economically attractive compared with no testing—no treatment if other features suggestive of Lyme disease were present. Other strategies were either more costly and less effective or were associated with incremental cost-effectiveness ratios of more than $50,000 per quality-adjusted life year.

Even if the analysis indicates a single best course of action, the differences in outcome and cost between the preferred strategy and the next best strategy should be large enough to be clinically meaningful. How large the difference must be to be important will depend on the outcome and the opinion of the reader. For life expectancy, some authors suggest a difference of 2 months or more is clinically meaningful. Unfortunately, there has not been a similar systematic review of differences in quality-adjusted life expectancy, but 2 months of relatively good health (utility weight of at least 0.8) would translate into 0.13 quality-adjusted life years. Other ad hoc outcomes (eg, number of strokes prevented) will be more subjective. In all cases it is important to remember that a decision analysis gives the expected average difference between groups and rarely is able (except when using the confidence interval technique above) to give a sense of the true variability around the average. If the variability is high, clinicians will see many patients whose outcomes are much better or much worse than the number stated in the model. In the Lyme disease model, the differences ranged from 0.003 to 0.019 quality-adjusted life years. This difference might not be considered clinically meaningful.

If one considers cost and outcome simultaneously (ie, the incremental cost-effectiveness of the alternative interventions), then different thresholds might apply. Some have proposed that $20,000 and $100,000 per quality-adjusted life year are reasonable boundaries for determining whether a technology is cost-effective (less than $20,000 per quality-adjusted life year [grade B]), moderately cost-effective ($20,000 to $100,000 per quality-adjusted life year [grade C]) or relatively cost-ineffective (greater than $100,000 per quality-adjusted life
year (grade D)). In this case, the recommended strategies for Lyme disease in each clinical scenario are cost-effective.

Finally, readers must decide whether the strategies recommended by the decision analysis are feasible to implement in their own clinical practice. Often, the decision problem has several implicit but unstated assumptions that are critical to a specific situation. For example, a colleague and I used decision analysis to determine whether computed tomographic (CT) scanning of the cervical spine would be preferred to the standard practice of plain radiographs for screening examinations of patients who come to the emergency department with head trauma. Although CT scanning was found to be effective and cost-effective, it is not a feasible alternative for those who work in emergency departments without ready access to CT scanners (unless safely transferring the patient to a facility with a CT scanner is an option). Decision analyses that involve cost-effectiveness often ignore the start-up costs of acquiring the technology of interest. If the start-up costs are great, the cost-effectiveness for the clinician who does not have ready access to the technology might be very different from what has been presented in the decision model.

In the case of the decision analysis for Lyme disease, you find that the results are clear and the magnitude of the difference between strategies compelling enough to warrant using this approach in your clinical practice. The laboratory where you practice is able to perform all the tests included in the decision analysis, so the strategy is feasible. Because the patient with oligoarthritis does not have other symptoms or signs suggestive of Lyme disease (history of tick bite or rash), the strategy of no testing and no empirical treatment is adopted.

Conclusions
Decision analysis can be a useful tool for guiding clinical practice policy and clinical decision-making. Decision analysis can help guide practice in situations where either direct evidence from clinical trials is unavailable or where the studies do not pertain directly to patient or patient group of interest. When reading a published decision analyses, it is important that clinicians consider how the study applies to their patient or patient group, whether the methods used to address the problem are appropriate, and whether the results provide information that can help them improve the care of a particular patient or patient group. Although decision analyses are not a substitute for high-quality clinical trials, their systematic approach to synthesizing the evidence can help clinicians make more informed choices about the care of their patients.

References
6. Richardson WS, Detsky AS. 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? Evidence Based Medicine Working Group. JAMA 1995;273:1610-3.
7. Richardson WS, Detsky AS. Users' guides to the medical literature. VII. How to use a clinical decision analysis. A. Are the results of the study valid? Evidence-Based Medicine Working Group. JAMA 1995;273:1292-5.


