Representation and Analysis of Medical Decision Problems with Influence Diagrams

DOUGLAS K. OWENS, MD, MSc, ROSS D. SHACHTER, PhD, ROBERT F. NEASE, Jr., PhD

Influence diagrams are a powerful graphic representation for decision models, complementary to decision trees. Influence diagrams and decision trees are different graphic representations for the same underlying mathematical model and operations. This article describes the elements of an influence diagram, and shows several familiar decision problems represented as decision trees and as influence diagrams. The authors also contrast the information highlighted in each graphic representation, demonstrate how to calculate the expected utilities of decision alternatives modeled with an influence diagram, provide an overview of the conceptual basis of the solution algorithms that have been developed for influence diagrams, discuss the strengths and limitations of influence diagrams relative to decision trees, and describe the mathematical operations that are used to evaluate both decision trees and influence diagrams. They use clinical examples to illustrate the mathematical operations of the influence-diagram–evaluation algorithm; these operations are arc reversal, chance node removal by averaging, and decision node removal by policy determination. Influence diagrams may be helpful when problems have a high degree of conditional independence, when large models are needed, when communication of the probabilistic relationships is important, or when the analysis requires extensive Bayesian updating. The choice of graphic representation should be governed by convenience, and will depend on the problem being analyzed, on the experience of the analyst, and on the background of the consumers of the analysis. *Key words*: decision analysis; influence diagrams; cost–effectiveness analysis; Bayesian updating; graphic representation.
(Med Decis Making 1997;17:241–262)

Decision models perform several functions in the analysis of medical problems. They enable clinicians and analysts to assess the expected utilities of alternative actions in situations that involve uncertainty, complexity, and dynamic change; to communicate explicitly assumptions about the structure of a problem; to determine the importance of uncertainty with sensitivity analyses; to determine the benefit of gathering further information through value-of-information calculations; and to make probabilistic inferences conditioned on evidence. The most familiar graphic approach for creating decision models is the decision tree. In this article, we describe an alternative graphic approach to analyzing decision problems: the influence diagram.

Received December 5, 1995, from the VA Palo Alto Health Care System, Palo Alto, California (DKO); the Section on Medical Informatics, Department of Medicine, and Department of Health Research and Policy (DKO) and the Department of Engineering–Economic Systems (RDS), Stanford University, Stanford, California; and the Laboratory for Medical Decision Sciences, Division of General Medical Sciences, Department of Internal Medicine, Washington University, St. Louis, Missouri (RFN Jr.). Revision accepted for publication August 14, 1996. Supported in part by the Department of Veterans Affairs Northeast Region Outcomes Project, and in part by grant DA 09531-01A2 from the National Institute on Drug Abuse to the Societal Institute of the Mathematical Sciences. Dr. Owens is supported by VA Health Services Research and Development Career Development Award. Dr. Nease is a Picker/Commonwealth Scholar.

Address correspondence and reprint requests to Dr. Owens: Section of General Medicine (111A), VA Palo Alto Health Care System, 3801 Miranda Avenue, Palo Alto, CA 94304.

A Graphic Representation of Decision Models

Influence diagrams, first used almost 20 years ago,[1] are graphic representations of formal mathematical models. They can be used to perform all the functions listed in the preceding paragraph, as can the decision-tree representation. So, if the two rep-

resentations perform similar functions, why use an influence diagram? Influence diagrams provide a complementary approach to the decision-tree representation. Influence diagrams represent the probabilistic structures of complex problems compactly, facilitate communication between analysts and content experts, and may have computational advantages for certain problems. These features have led investigators to use influence diagrams for Bayesian meta-analysis,[2,3] for representing bias in clinical studies,[4,5] for developing patient-specific explanations,[6] for developing clinical guidelines,[7] and to explain probabilistic inferences.[8] Influence diagrams also have limitations relative to decision trees. An important key to understanding influence diagrams is the realization that influence diagrams and decision trees are simply different graphic representations for the same underlying mathematical model and operations.

In this article, we explain how problems are analyzed in influence diagrams. In a companion article,[9] we show how to structure problems as influence diagrams. The next section introduces the terminology and notation of influence diagrams. In the following section, familiar decision problems are represented as decision trees and as influence diagrams, and arc reversal, an important operation for the evaluation of influence diagrams, is introduced. The next section describes an algorithm to evaluate influence diagrams, and provides several examples. The final section discusses the advantages and limitations of influence diagrams relative to decision trees. The appendix further describes the mathematical operations used in the evaluation of influence diagrams.

Definitions and Notation

Influence diagrams represent graphically many of the same components of a decision model as do decision trees. Figure 1 shows a decision tree and the corresponding influence diagram. The decision problem modeled by the tree and by the influence diagram is the decision to test an infant born to a mother who is infected with the human immunodeficiency virus (HIV) with the polymerase chain reaction (PCR), a gene-amplification technique that is useful for diagnosis of HIV infection in infants. Both the decision tree and the influence diagram show two decisions. The first decision is whether to perform PCR, and the second is whether to treat [begin prophylaxis for infection with *Pneumocystis pneumoniae* (PCP)] once the test result is known. Current guidelines recommend initiation of PCP prophylaxis in infants infected with HIV.[10] The benefit modeled is that of a longer quality-adjusted life expectancy if PCP prophylaxis is used in an HIV-infected infant. The outcome of interest is the quality-adjusted length of life. Two new graphic elements are apparent in the influence diagram: arcs between nodes, and the value node (shown as a diamond).

The *arcs* represent relationships between the *nodes*. A *decision node* (drawn as a square) provides the decision alternatives under consideration. A *chance node* (drawn as a circle) represents a variable whose value is a probabilistic function. An arc between two chance nodes indicates that a probabilistic relationship between the two events *might* exist. A probabilistic relationship exists when the occurrence of one of the events affects the probability of the occurrence of the other event. We know, for example, that the PCR test result (the chance node PCR Result in figure 1) depends on whether the infant is infected (HIV Status, figure 1). The arc between HIV Status and PCR Result indicates this probabilistic dependence. An arc represents a *weak* assertion about a probabilistic relationship, because an arc is allowable between two chance nodes when, in fact, no probabilistic relationship exists.

The arc points from the conditioning event to the conditioned event. Thus, in figure 1, the arc from HIV Status to PCR Result indicates that the test result is conditioned on the infection status; the diagram requires an assessment of the probability of a positive or negative test conditioned on whether the infant is infected. The direction of the arc, therefore, determines which probabilities will be assessed as conditional probabilities. As we shall show, the direction of arcs may be changed with Bayes' theorem during evaluation of an influence diagram. Therefore, the direction of the arc *does not* imply causality.

The *absence* of an arc is a *strong* assertion of independence or of *conditional independence*. Two events are conditionally independent, given a third event, if, after we have observed the third event, observing one of the two events gives us no additional information about the likelihood of the other event. In figure 1, there are arcs from HIV Status and Treat? to QALE (quality-adjusted life expectancy), but no arc from PCR Result to QALE. This absence of an arc indicates that QALE is conditionally independent of PCR Result, given knowledge of HIV-infection status and whether the patient was treated (Treat?). QALE is conditionally independent of PCR Result because, if we know whether the patient truly has HIV infection and whether he or she was treated, information about the test result, per se, does not affect quality-adjusted life expectancy. Arcs between chance nodes can be omitted only when events are assumed to be conditionally independent.

In contrast to a decision tree, in which the sequence of events is evident from the tree structure, an influence diagram relies on specific types of arcs

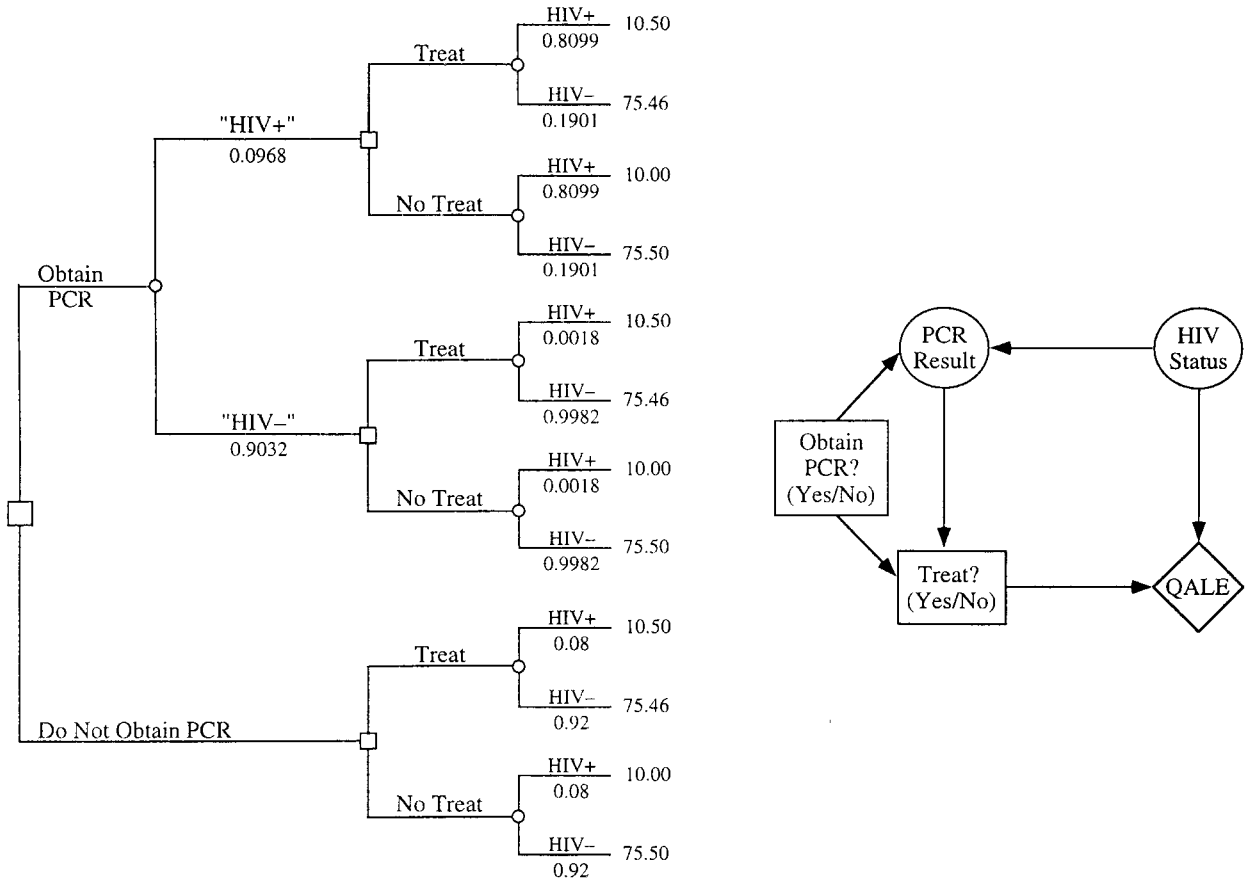


FIGURE 1. Alternative graphic representations of the decisions to test for and to treat human immunodeficiency virus (HIV) infection. The decision tree (left) shows the structural asymmetry of the problem; the influence diagram (right) displays the informational and probabilistic relationships. HIV = human immunodeficiency virus; HIV+ = HIV-infected; HIV– = not infected with HIV; QALE = quality-adjusted life expectancy; PCR = polymerase chain reaction. Test results are shown in quotation marks ("HIV+") and the true disease state is shown without quotation marks (HIV+).

to represent the sequence of events. An arc that points into a decision node from a chance node indicates that the chance event has been observed (or is known) at the time the decision is made. Thus, in figure 1, the arc from PCR Result to Treat? indicates that the decision maker knows the PCR result prior to making the decision to treat, if the decision maker obtained a PCR test. Such arcs are called *informational arcs.* Conversely, the absence of an arc from a chance node (HIV Status) to a decision (Treat?) indicates that the decision maker has not observed the outcomes of the chance event when he or she makes the decision. Thus, figure 1 asserts that the decision maker does not know the actual HIV-infection status when he or she decides whether to treat. An arc that points from a decision node A (Obtain PCR?) into decision node B (Treat?) indicates that decision A is made prior to decision B. These arcs are often called "*no-forgetting arcs,*" indicating that the decision maker does not forget decisions that were made previously, or forget the information available at the time of the earlier decisions. (A decision tree requires these assumptions as well.) The analyst must specify completely the order of decisions in the influence diagram; the no-forgetting arcs enable the analyst to indicate this ordering. In figure 1, for example, the no-forgetting arc from Obtain PCR? to Treat? indicates that the decision maker decides whether to obtain PCR prior to deciding whether to initiate PCP prophylaxis (Treat?) and that the decision maker remembers whether he or she obtained a PCR test at the time of deciding whether to treat.

The information associated with a node is determined by the node type and by the node's parents. The *parents* (or *direct predecessors*) of a node are those nodes that send arcs to the node; the *children* (or *direct successors*) of a node are those nodes that receive arcs from the node. A chance node provides the probability of the outcome of events conditioned on the node's parents. For example, in figure 1, PCR Result provides the probability of a positive test and the probability of a negative test conditioned on HIV Status and on whether a test was ordered. A *deterministic node* is a special chance node that represents a variable whose value is a deterministic func-

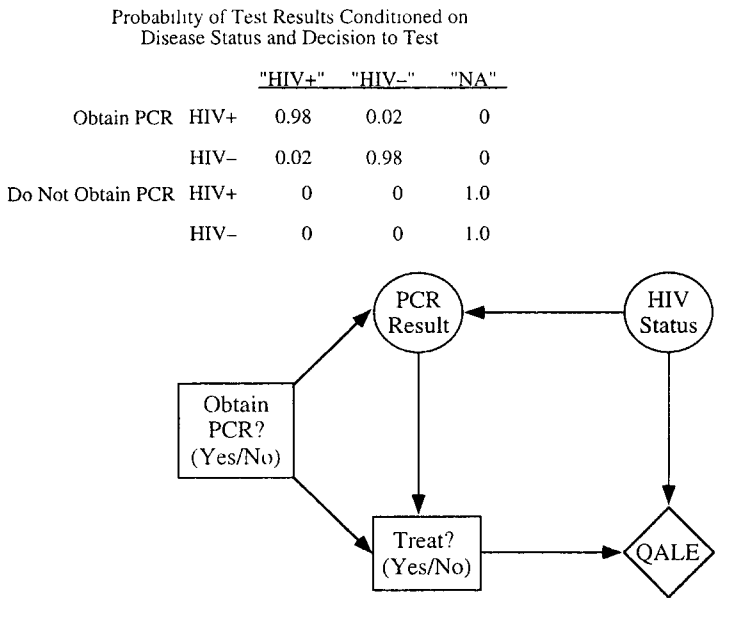


FIGURE 2. Probability and value tables associated with an influence diagram. The information that is associated with branches and endpoints of a decision tree is placed in tables in the influence diagram. HIV = human immunodeficiency virus; HIV+ = HIV-infected; HIV− = not infected with HIV; QALE = quality-adjusted life expectancy; PCR = polymerase chain reaction; Tx+ = treated; Tx− = untreated. Test results are shown in quotation marks ("HIV+") and the true disease state is shown without quotation marks (HIV+).

tion of its parents. A deterministic function differs from a probabilistic function in that if the value of the parents of a deterministic node are known, the value of the deterministic function is known with certainty. Deterministic nodes usually are drawn as a double circle in influence diagrams, to distinguish them from ordinary chance nodes. The information associated with a node is contained in a table, as we shall explain.

The node labeled QALE in figure 1 is a *value node*, drawn as a diamond. The value node contains the information that is shown at the ends of the branches of a decision tree. The parents of the value node (i.e., nodes that have arcs that point to the value node) indicate the events and decisions that affect the utility directly. The parents of the value node depend on the utility model used for the analysis.[9] In figure 1, we see that only HIV Status and Treat? affect value (quality-adjusted length of life) directly.

The value node and decision nodes distinguish an influence diagram from a *belief network*, which is an influence diagram that contains only chance nodes.[11] Belief networks are used to perform probabilistic inference, but do not allow for the evaluation of decisions or for representation of utility models. Belief networks are useful for large probabilistic inference problems, such as those in medical diagnosis.[12,13]

Figure 2 shows the influence diagram from figure 1, with tables adjacent to the individual nodes. These tables contain the information that would be associated with the branches in a decision tree. Two types of tables are shown in figure 2: probability tables and a value table. The probability tables contain the probabilistic information that relates conditioned events and conditioning events. For example, the probability table associated with the PCR Result node in figure 2 denotes the probabilistic relationship of the outcomes of the PCR Result chance node to the conditioning events, which are the decision Obtain PCR? and HIV Status. The notations that we use in the figures and text are defined in table 1. Our convention for these tables is that we place the conditioning events (Obtain PCR, HIV+) in a column at the left of the table, and place the possible values for the variable ("HIV+", "HIV−", "NA") across the top of the table. The probabilities in each row of the table must sum to 1.0, because the row contains all possible outcomes conditioned on the events in the first column. Notice that we indicate a test result with quotation marks, and the true disease state without quotation marks. The probability table associated with PCR Result indicates that, if the decision alternative Obtain PCR is chosen and if the patient is infected with HIV (HIV+), the probability of a positive test result ("HIV+") is 0.98,[14] and the probability of a negative test result ("HIV−") is 0.02. The table indicates that, if the decision alternative Do Not Obtain PCR is chosen, then the test result is "not available" ("NA") with probability 1.0 (right column), regardless of the infection status of the patient. The need to specify probabilities for the test result when the Do Not Obtain PCR alternative is chosen arises because the influence diagram assumes symmetry of the alternatives and uncertain events. When a decision problem has different events depending on the actions taken (for example, the events following the Obtain PCR and Do Not Obtain PCR alternatives are different), we say that the problem has *structural asymmetry*. Thus, in the influence diagram, asymmetry in the structure of a decision is denoted by outcomes that have a probability of 1 or 0.

The value table is best understood in relation to

the value node. The arcs into the value node (from the parents) indicate the events that affect the values of the outcomes. The value table is associated with the value node; as noted, it reflects how the decision maker values the possible outcomes that may result from the decision. In figure 1, for example, the outcomes in the decision tree and the influence diagram are valued in terms of quality-adjusted life expectancy. The arcs into the value node indicate that the quality-adjusted life expectancy depends on whether the screened person is infected (HIV Status, in figure 1), and on whether treatment (PCP prophylaxis) is initiated. The value table indicates that an initiation of PCP prophylaxis increases quality-adjusted length of life by six months (in fact, the degree of benefit is uncertain, but we assume six months of benefit) for infants infected with HIV. We assume that adverse effects from pharmacologic PCP prophylaxis result in a decrease of quality-adjusted life expectancy of approximately 0.04 years for uninfected infants. Thus, in the influence diagram, the values of the outcomes are placed in the value table; in a decision tree, they would be placed at the ends of the branches.

Examination of figures 1 and 2 reveals several differences between the information displayed in the decision tree and that displayed in the influence diagram. First, the decision alternatives, the outcomes of the chance events, and the probabilities associated with these outcomes are not shown graphically in the influence diagram. In the influence diagram, this information is contained in tables associated with the corresponding nodes (fig. 2). In software developed to analyze influence diagrams, data entry is usually made directly into such tables. The emphasis on the probabilistic structure of the problems in the graphic representations of influence diagrams enhances the capability for displaying the probabilistic relationships in large, complex problems; however, the probabilities and utilities of outcomes are not apparent in the graphic representations.

Notice that the decision tree shows structural asymmetry (for example, the Obtain PCR branch and the Do Not Obtain PCR branch of the decision tree in figure 1 are different) graphically, and that the influence diagrams highlight the probabilistic relationships in the model. For example, the influence diagram indicates that quality-adjusted life expectancy is independent of the test result, given knowledge of the infection status and of whether PCP prophylaxis was initiated. This probabilistic independence is not explicit in the decision tree. As noted, the absence of an arc from Obtain PCR? to QALE indicates that, conditioned on the disease state and treatment decision, the test is assumed to have no effect on utility (for example, a harmless test). Assumptions about probabilistic independence are explicit in the influence diagram, even in extremely complex models. Such assumptions are more difficult to identify with a decision tree: the analyst must compare the probabilities on all branches in the tree for the relevant events. A decision tree shows structural asymmetry in branches that model different events and, thus, differ in their structures. In the influence diagram, however, structural asymmetry is hidden; the analyst indicates asymmetry by assigning a probability of 0 to certain events. For problems with considerable structural asymmetry, the decision tree may be a more natural representation.

Table 1 • Notations

Symbol	Meaning
A+	Event A occurs
A−	Event A does not occur
A	Event A+ or A−
p(A+)	Probability that event A occurs
p(A+|B+)	Probability that event A occurs given that event B occurred
Tx+	Treatment administered (a decision alternative)
Tx−	Treatment withheld (a decision alternative)
HIV+	Disease present (true state, HIV = human immunodeficiency virus)
HIV−	Disease absent (true state)
"HIV+"	Test result indicates that disease is present
"HIV−"	Test result indicates that disease is absent
U[Tx+, HIV+]	Utility when treatment is offered and disease is present
EU|"HIV+", Tx+	Expected utility when the test result indicates disease and treatment is offered

Influence Diagrams and Trees: Examples

In this section, we provide examples of familiar medical decision problems, represented both as influence diagrams and as decision trees. The tree and influence diagram in figure 3a show a treatment decision for HIV infection in the infant when test information is not available. From the influence diagram, we see that the value depends on whether treatment is offered and on whether HIV infection is present. The tree and influence diagram in figure 3b represent a treatment decision in which test information is available (PCR Result, as shown in bold). Our decision tree begins with a chance node, because we assume that the decision maker has already ordered a test, that he or she has not yet received the result (therefore, the result is uncertain), and that he or she will have observed the test result when making the decision to treat. From the influence diagram, we notice again that the value de-

pends on whether treatment is offered and on whether disease is present. The PCR Result node is probabilistically related to the HIV Status node, as indicated by the arc from HIV Status to PCR Result. The informational arc from PCR Result to the Treat? decision node indicates that the decision maker will have observed the test result at the time that he or she makes the decision. The absence of an arc between HIV Status and Treat? indicates that the decision maker does not know the actual disease status at the time of making the decision (based on the test result), and that the decision does not affect Disease Status.

For the simple problems shown in figures 1 and 3, the decision-tree representation is appealing. For large and complex problems, however, representation in a tree can become more challenging. The use of subtrees is one familiar and effective approach to displaying large models; influence diagrams also provide a compact representation. Both subtrees and influence diagrams represent decision problems with structural symmetry more easily than they represent asymmetric problems. In figure 4, we show an influence diagram for a model of a recently published cost–effectiveness analysis of a program to screen surgeons for HIV infection to prevent transmission to patients.[15,16] To understand the influence diagram, we can work back from the value node on the right side of the figure. Notice first that the value node is affected directly by QALE. Examine the parents of QALE and notice that we modeled three benefits of a screening program: the benefit from reduced transmission to patients during surgery (from a policy that restricts surgeons identified as being infected with HIV from performing procedures), the benefit from reduced transmission of HIV to the surgeons' sexual partners, and the benefit to the surgeon screened and identified as infected (from early medical intervention). The degrees to which these benefits are realized depend in turn on

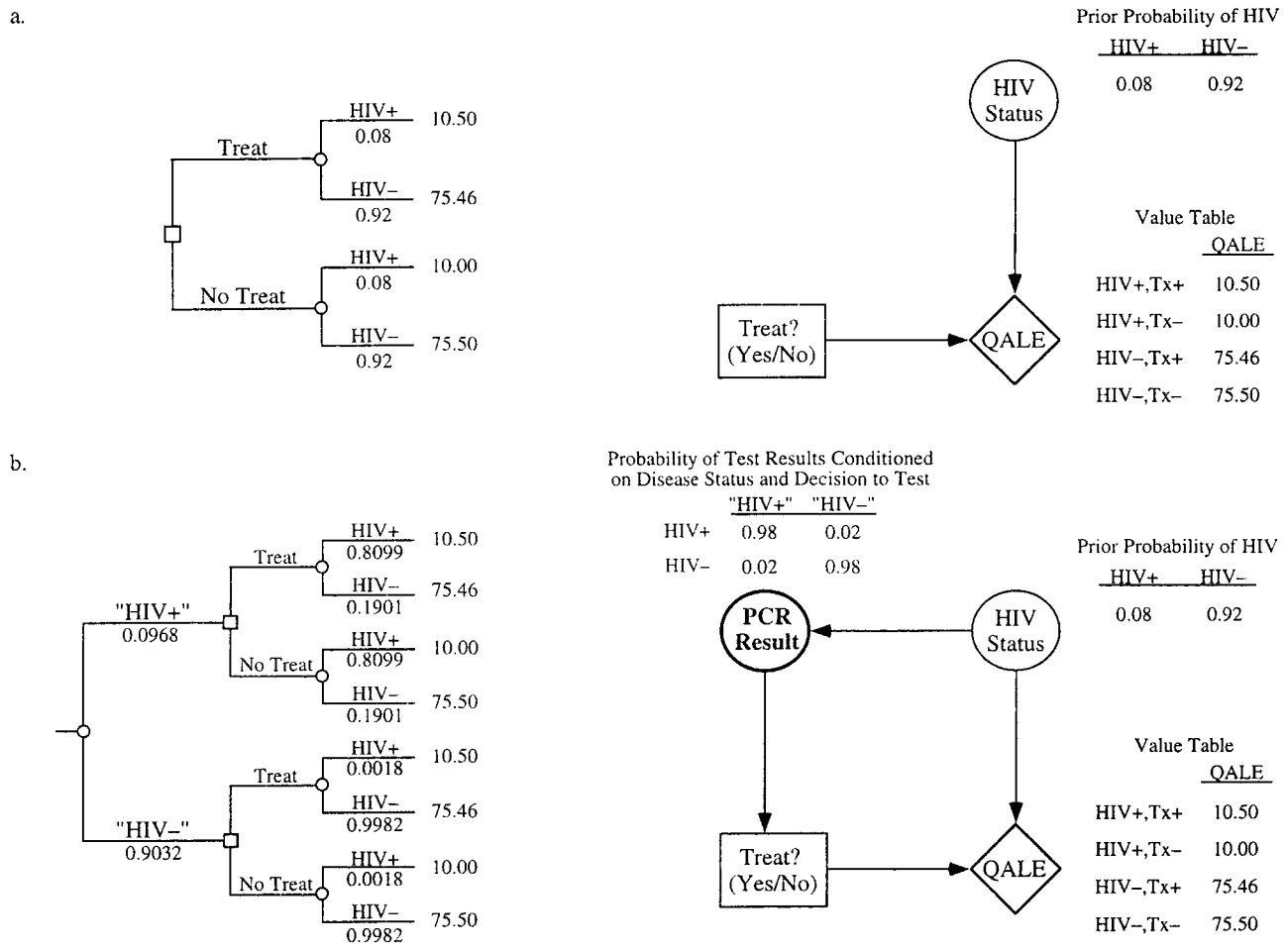


FIGURE 3. Example decision problems. In each panel, a decision problem is represented graphically as both a decision tree and an influence diagram. In (a), treatment is considered for HIV; test information is not available. In (b), a test has been ordered already and the result will be known at the time the Treat? decision is made. The decision tree is shown in order of observation with posterior probabilities of HIV infection conditioned on the test result. The influence diagram is shown in assessment order. HIV = human immunodeficiency virus; HIV+ = HIV-infected; HIV– = not infected with HIV; QALE = quality-adjusted life expectancy; PCR = polymerase chain reaction; Tx+ = treated; Tx– = untreated. Test results are shown in quotation marks ("HIV+") and the true disease state is shown without quotation marks (HIV+).

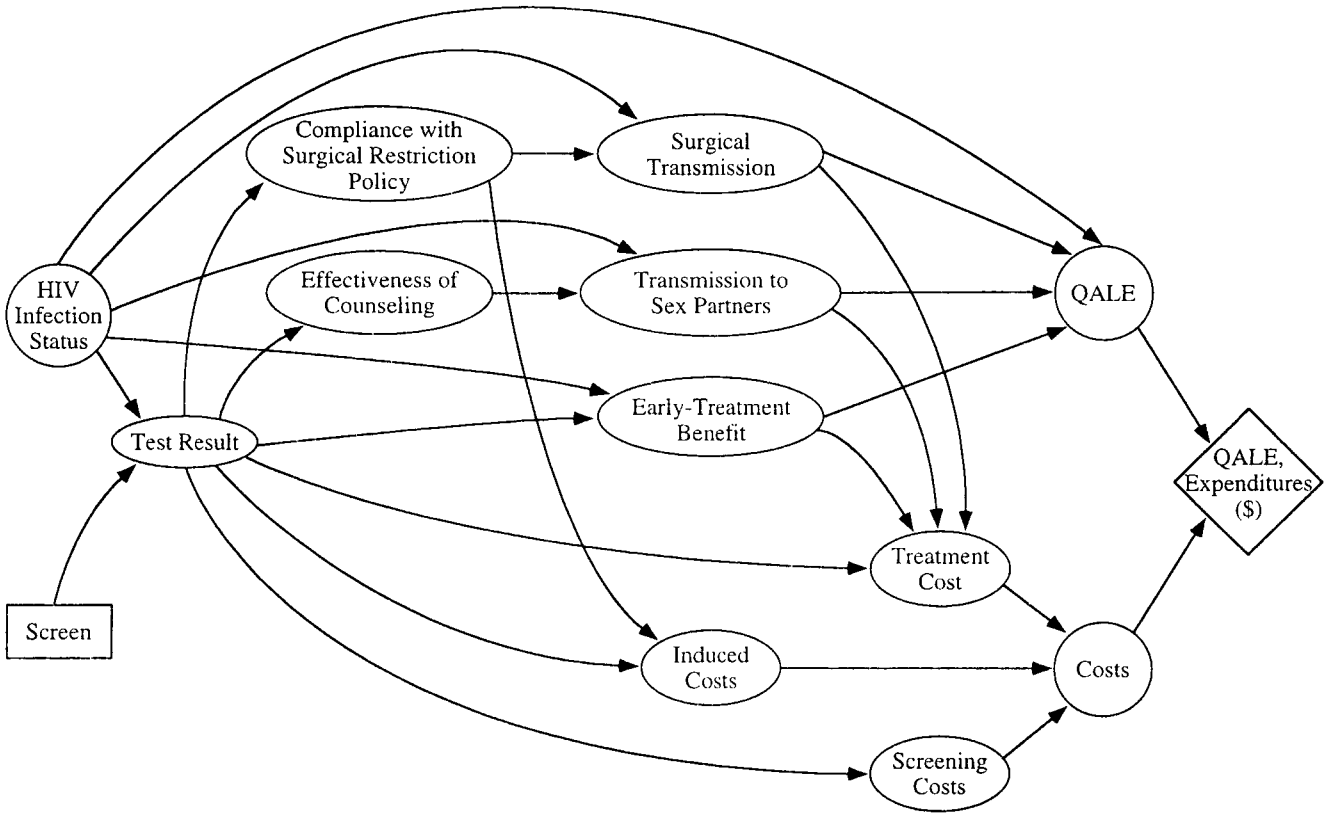


FIGURE 4. Screening surgeons for human immunodeficiency virus (HIV) infection. The influence diagram shows a model of the decision to screen surgeons for HIV infection and to prohibit from operating those surgeons who are infected. See text. HIV = human immunodeficiency virus; QALE = quality-adjusted life expectancy.

compliance with the policy to restrict surgical practice and on the effectiveness of counseling in inducing reductions in sexual-risk behaviors. The influence diagram shows that each of these benefits depends on the test result, which depends on infection status, and the decision to screen. Although we show in figure 4 that costs can be modeled in a similar manner, use of influence diagrams for cost–effectiveness analysis is beyond our scope, and thus we do not discuss costs further here. Conditional independence is shown by the absence of arcs.

CONVERSION BETWEEN GRAPHIC REPRESENTATIONS

Because decision trees and influence diagrams are differential graphic representations of similar underlying mathematical operations, it is possible to convert one graphic representation into the other. Any symmetric decision tree can be drawn as an influence diagram without further manipulation; however, all influence diagrams cannot be drawn as decision trees directly. Decision trees usually are drawn from left to right, with events in the *order of observation relative to the decisions.* Thus, an event that is observed by the decision maker before he or she makes a decision is drawn to the left of the decision; an event that is observed after the decision is drawn to the right of the decision. (An alternative, however, is for the analyst to put all decisions at the left of the tree—a representation called normal form.[17]) To draw the tree in order of observation, the analyst may need to use Bayes' theorem to calculate the probabilities required in the tree (the probabilities of a positive and of a negative test result, and the posttest probability of disease given each test result). Because an influence diagram does not require events to be drawn in the order of observation, the analyst may need to manipulate the diagram before converting it to a decision tree that is drawn in the order of observation. This manipulation is mathematically identical to the use of Bayes' theorem to calculate probabilities required in the decision tree. Applying Bayes' theorem in an influence diagram involves changing the direction of the arcs, an operation called *arc reversal.* Although a full discussion of how to convert influence diagrams to trees is beyond the scope of this paper (for a detailed explanation, see Schachter[18] and Howard and Matheson[19]), we now review arc reversal, because it is an essential operation for the evaluation of an influence diagram.

ARC REVERSAL

An influence diagram can be drawn to reflect either of two different orders of probabilistic conditioning. We can draw the diagram with variables ordered according to when they will be observed (as in a decision tree). One of the advantages of the influence diagram, however, is that it can be drawn with the variables ordered such that the probabilities can be assessed most easily. This order of conditioning is called the *assessment ordering*. If, for example, we have evidence about the prevalence of disease and about the sensitivity and specificity of the diagnostic test, the assessment ordering is disease status (for example, HIV Status, figure 1), followed by the test result (PCR Result, figure 1). The ability to use assessment ordering freely can ease the analytic task substantially. It allows the analyst to build the diagram with events in the order that most facilitates probability assessment [for example, from cause (disease status) to effect (test result)[20]], an approach to structuring the problem that is more natural than the reverse ordering.[9]

We can illustrate the concepts underlying arc reversal with an example. Suppose that we want to determine the probability that a person who has a history of injection drug use and needle sharing is infected with HIV. Let's assume that two tests are available: PCR, which has a sensitivity and specificity of 98.1 in adults (PCR sensitivity and specificity need not be identical, but recent evidence indicates that sensitivity = specificity = 98.1 represents the upper left point on a receiver operating characteristic curve),[21] and the standard antibody tests, which have a sensitivity of 99.5% or greater, and a specificity of 99.99%.[16] The influence diagram in figure 5a illustrates what we know. We believe that HIV Status depends on Risk Behavior—in this case, needle sharing. More specifically, the prior probability of HIV infection depends on risk behavior. Test Result depends on Type of HIV Test (because one test is more accurate than the other, and thus the sensitivity and specificity will differ), and on HIV Status. Because Test Result is conditioned on HIV Status (rather than HIV Status being conditioned on Test Result), the diagram is drawn in assessment ordering, and we can enter the sensitivity and specificity of each test directly in the probability table associated with the node Test Result. The diagram asserts that Test Result is independent of Risk Behavior, given that we know HIV Status. This independence is indicated by the absence of an arc from Risk Behavior to Test Result, and indicates that if we knew the true HIV-infection status of an individual, information about the individual's risk behaviors would not affect the probability of a positive or negative test result. We also notice from the diagram that Risk Behavior and HIV Status are independent of Type of HIV Test, as shown by the absence of arcs between these nodes. We could complete the probability tables in figure 5a with information that we already know—the prevalence of HIV infection conditioned on risk behavior (for HIV Status) and the sensitivity and specificity of the HIV tests (for Test Result).

The information we seek from the influence diagram, however, is the posttest probability of HIV infection, given a particular test result. That is, what is HIV Status, given Test Result? We can answer this question by reversing the arc between HIV Status and Test Result, as shown in figure 5b; this diagram indicates that HIV Status is conditioned on Test Result. The probability table associated with HIV Status would therefore contain the probability of HIV infection conditioned on a positive or negative test (we show this calculation in detail later). But is figure 5b otherwise correct? It asserts that HIV Status depends on Test Result only (that is, whether the test result is positive or negative), and is independent of Type of HIV Test. We know, however, that the posterior probability of HIV infection depends on which test we choose, because the tests have different sensitivities and specificities. Thus, simply reversing the arc between HIV Status and Test Result creates a diagram that incorrectly represents our knowledge of the problem. To correct this problem, we must first add an arc from Type of HIV Test to HIV Status, as shown in figure 5c. This new arc reflects the dependence of the posterior probability of HIV infection on the type of test that we choose. Is figure 5c correct? It asserts that Test Result depends on Type of HIV Test only, and is independent of Risk Behavior. However, we know that Test Result depends on Risk Behavior; a positive test result is more likely among those patients who engage in risky behavior. To reflect this dependence, we must add another arc from Risk Behavior to Test Result, as shown in figure 5d. Notice that now both HIV Status and Test Result are conditioned on the same events. We now have a diagram that represents our knowledge faithfully, and because HIV Status is conditioned on Test Result, we should be able to determine the posttest probability of HIV infection.

Our reasoning about this example illustrates a general principle. To reverse an arc between node A and node B, we must ensure that node A and node B have the same parents—A and B must be conditioned on the same events. Often, as we did in figure 5, we must add arcs to the diagram to meet this criterion. The reason is that mathematically, we perform arc reversal using Bayes' theorem. Bayes' theorem requires that events A and B be conditioned on the same events. (Recall that adding an arc means not that a probabilistic relationship does exist, but rather that a probabilistic relationship could

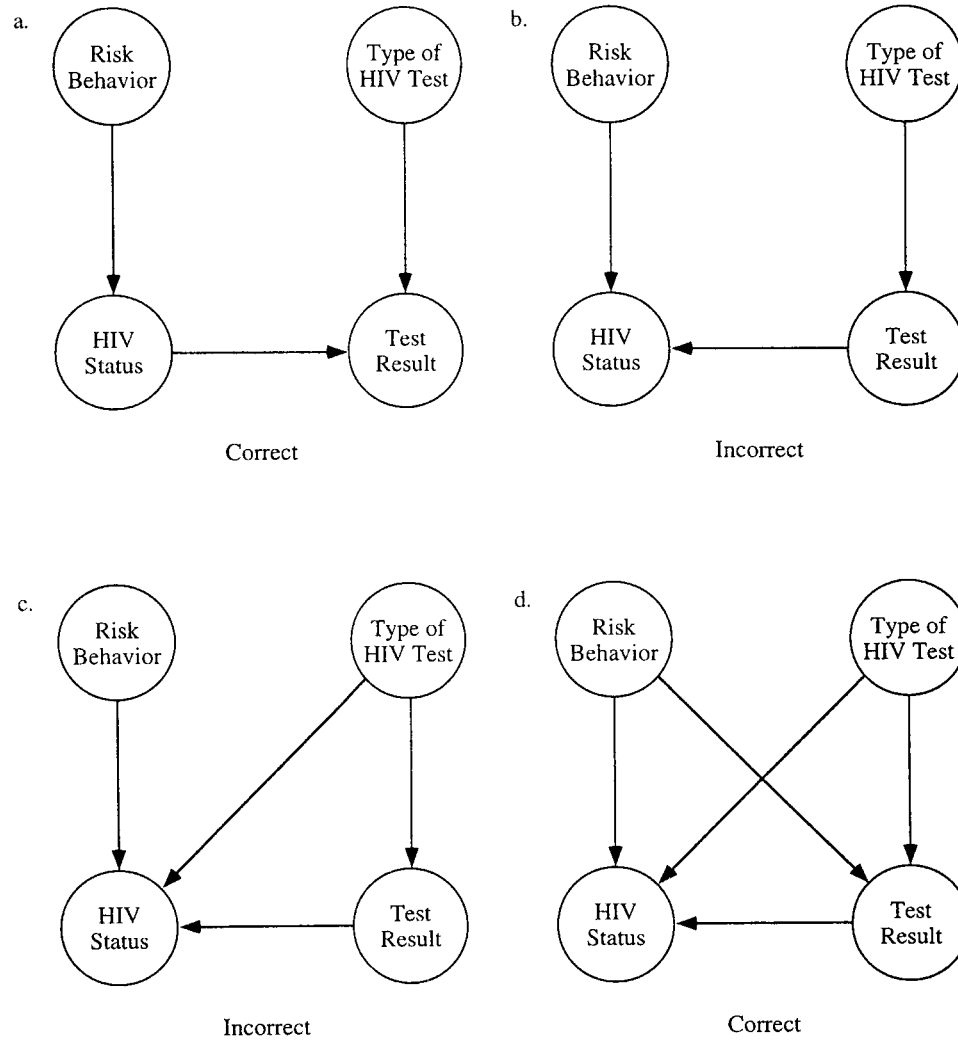


FIGURE 5. Arc reversal. To reverse the arc between HIV Status and Test Result, additional arcs must be added to ensure that these nodes have identical parents. HIV = human immunodeficiency virus.

exist.) In summary, we can draw the influence diagram such that we facilitate data entry, and we can use arc reversal to perform inference. We now turn to a method for evaluating influence diagrams more completely, so that we can use influence diagrams to evaluate decision alternatives.

Evaluation of Influence Diagrams

The mathematical operations used to evaluate influence diagrams and decision trees are the operations of probability and expected-utility theory. The key concept is that decision trees and influence diagrams provide different graphic representations of the same underlying probability distributions and expected utility operations. The procedure for evaluating the decision tree is familiar and can be performed easily by hand. Starting at the right side of the tree, the analyst takes the expectation at each chance node (by adding the products of the probability and the utility for each branch), substitutes the expectation for the chance node, and repeats the process until he or she reaches a decision node. Thus, the analyst removes the nodes by taking the weighted average of the branches of the node. We call this process *chance-node removal by averaging* (or *by integration*, for the continuous case). At the decision node, the analyst chooses the alternative with the highest expected utility, and then removes the decision node. We call this process *decision-node removal by policy determination.* Notice that the tree is successively whittled away and simplified as it is evaluated. This process continues until the analyst can determine the expected utilities of the decision alternatives at the leftmost decision.

Evaluation of an influence diagram is similar—the analyst simplifies the diagram successively until he or she can evaluate expected utilities of the decision alternatives. The process for simplifying the diagram also makes use of chance-node removal by averaging and decision-node removal by policy determination. Because the analyst may have drawn the influence diagram in assessment ordering, he or she may need to reverse arcs (and therefore to add arcs) before removing a node by averaging.

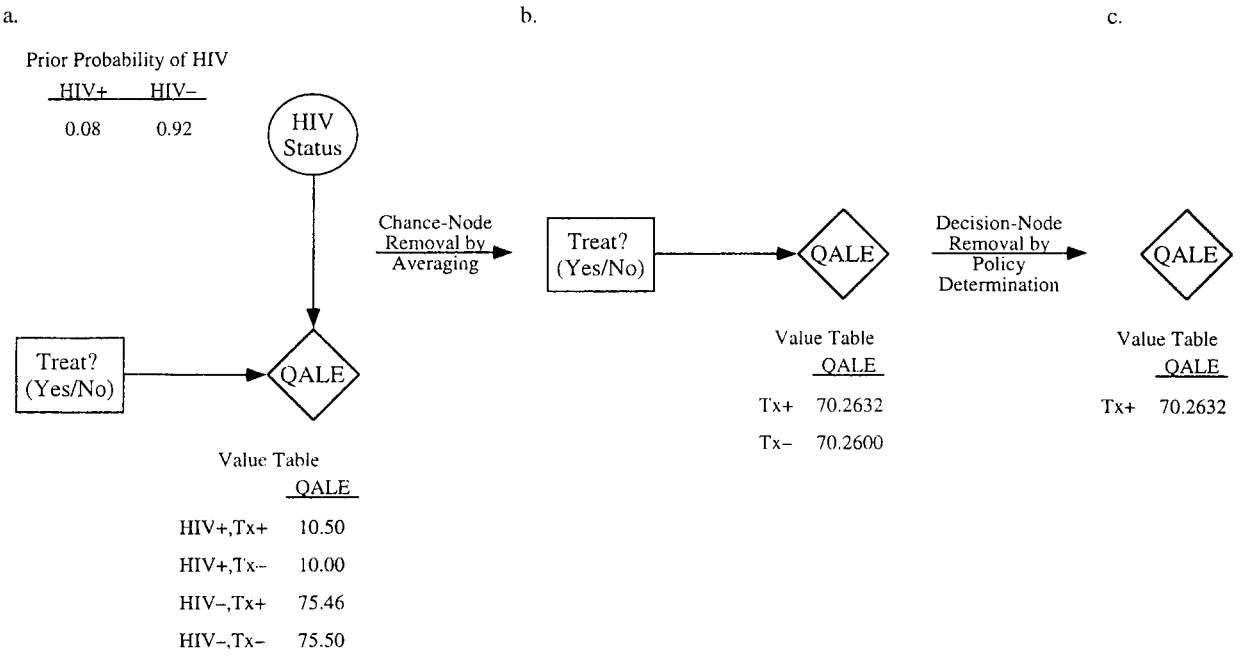


FIGURE 6. The Treat or No Treat decision. To evaluate the diagram, HIV Status (a) is removed by averaging (b), then Treat? is removed by policy determination (c). HIV = human immunodeficiency virus; HIV+ = HIV-infected; HIV− = not infected with HIV; QALE = quality-adjusted life expectancy; Tx+ = treated; Tx− = untreated.

Investigators have developed several algorithms for evaluating influence diagrams. The simplest is the method of arc reversal and node removal, developed by Shachter,[22,23] which we present below. To gain additional efficiency when there are many arcs to reverse, certain algorithms incorporate techniques used to perform probabilistic inference based on undirected graphs.[24–28] The undirected-graph algorithms are of the same computational complexity as is the influence diagram arc reversal/node removal algorithm, but they can be implemented more efficiently.[20,29] The undirected-graph techniques have been incorporated into several fundamentally similar algorithms for the evaluation of influence diagrams (or of related graphic constructs).[30–32] We discuss the mathematical basis for these algorithms further in the appendix. Although small influence diagrams can be evaluated by hand, in practice, software is required, much as it is for complex decision trees or for those trees that require substantial Bayesian updating. For both influence diagrams and trees, software also permits the analyst to perform sensitivity analyses easily.

EVALUATION ALGORITHM

We must remove nodes from an influence diagram in a specific order, just as we must evaluate the nodes in a decision tree in a specific order. The general principles are similar to those that are used to evaluate a decision tree. We first remove chance nodes for events whose outcomes are revealed, if ever, subsequent to a decision—these are events that the decision maker has not observed at the time of the decision. This step is analogous to folding back a decision tree from the right to the left. We remove decision nodes in reverse of the order in which we will actually make the decisions (as with a decision tree). That is, we remove the decision node for the final decision first, in a manner analogous to the evaluation of a decision tree that has sequential decisions. After removing the decision node for the final decision, we remove chance nodes whose events are revealed subsequent to the next-to-final decision that we must make, and so on. (Below, we illustrate these concepts with several examples.) We can evaluate influence diagrams with the following algorithm[22] that formalizes these concepts. No Bayesian updating is required prior to evaluation of the influence diagram with this algorithm.

1. Eliminate all nodes (except the value node) that do not point to another node (*barren nodes*). In the examples that we evaluate here, there are no such nodes. They may occur, however, when several types of evidence are observed in large models.

2. As long as there are one or more nodes that point into the value node, do the following:

   a. If there is a decision node that points into the value node, and if all other nodes that point into the value node also point into that

decision node, remove the decision node by policy determination. Remove any nodes (other than the value node) that no longer point to any other node. Go back to step 2.

b. If there is a chance node that points into only the value node, remove it by averaging. Go back to step 2.

c. Find a chance node that points into the value node and not into any decision. Reverse all of the arcs that point from that chance node into other chance nodes without creating a cycle (see the appendix). Now the chance node will point into only the value node (and therefore, will meet the criterion of step 2b). Go back to step 2.

In step 2a, we recognize that, if all chance nodes point into a decision node (that is, there are informational arcs from all chance events into a decision), then we have observed the outcomes of all chance events, and thus, we have no uncertainty about the relative ranking of the decision alternatives. We can, therefore, choose the best alternative, and remove the decision node by policy determination. In step 2b, we recognize that, if a chance node points into the value node and into no other decision nodes, the outcomes of the uncertain event are revealed subsequent to any decisions, so we can remove the chance node by averaging. Removing nodes by averaging is analogous to folding back the tree from the endpoints of the branches in the tree. In step 2c, we recognize chance nodes whose events are revealed subsequent to all decisions (these nodes will not have informational arcs into any decision node, which indicates that their outcomes are not observed prior to the decisions), but that must undergo arc reversal prior to removal. These are variables whose order we must reverse if we are perform the evaluation. The conditions addressed by step 2c occur only when we have entered probabilistic information into the influence diagram such that the application of Bayes' theorem is required during evaluation. The operation of arc reversal is analogous to the use of Bayes' theorem, which we often use in the process of building a decision tree (prior to evaluation of the tree). *As we remove a node of any type, we draw the arcs from its parents to its child, the value node.*

THE TREAT OR NO-TREAT DECISION

We now use the algorithm to evaluate several influence diagrams. We start with the problem in which the decision maker must decide whether to treat or not to treat, with no test information available (fig. 6a). There are no barren nodes (step 1). Because the conditions of step 2a are not satisfied —HIV Status does not point to the decision Treat?, which indicates that the decision maker has not observed HIV Status at the time that he or she makes the decision—the decision node cannot yet be removed. The conditions of step 2b are satisfied, however, so we remove HIV Status by averaging. To determine what calculation to perform in removing a node, we redraw the diagram with the arcs from the parents of the removed node pointing to its children, then examine the diagram to determine the appropriate conditioning for the new probability or value tables. If the node we wish to remove has no parent, as is the case in figure 6a, we examine the diagram to determine the appropriate conditioning, and then simply remove the node. We see from figure 6b that after we remove HIV Status, QALE will be conditioned on only Treat?. Therefore, we must calculate the expected utility for each of the decision alternatives for the Treat? decision using the information in the probability table associated with HIV Status and in the value table associated with QALE. Notice that performing this operation is identical to folding back the decision tree in figure 3a one level and removing the chance node from the tree. We calculate the expected utility for the treatment option (Tx+) first:

$$\begin{aligned} &EU[Tx+] \\ &= p(HIV+)U[HIV+, Tx+] + p(HIV-)U[HIV-, Tx+] \\ &= (0.08)(10.50\ QALYs) + (0.92)(75.46\ QALYs) \\ &= 70.2632\ QALYs \end{aligned}$$

The expected utility for the no-treatment option (Tx−) is

$$\begin{aligned} &EU[Tx-] \\ &= p(HIV+)U[HIV+, Tx-] + p(HIV-)U[HIV-, Tx-] \\ &= (0.08)(10.00\ QALYs) + (0.92)(75.50\ QALYs) \\ &= 70.2600\ QALYs \end{aligned}$$

We now have accounted for our uncertainty about HIV status, and have calculated the expected utility associated with each of our decision alternatives; we can therefore remove the node HIV Status (fig. 6b). Readers who wish to reproduce the numbers that we show in the figures should carry four decimals throughout all calculations to avoid errors from rounding off (the numbers we show agree with an analysis performed with software to at least the second decimal place).

Continuing to follow the algorithm, we return to step 2. We have removed by averaging all the chance nodes whose events are revealed subsequent to the

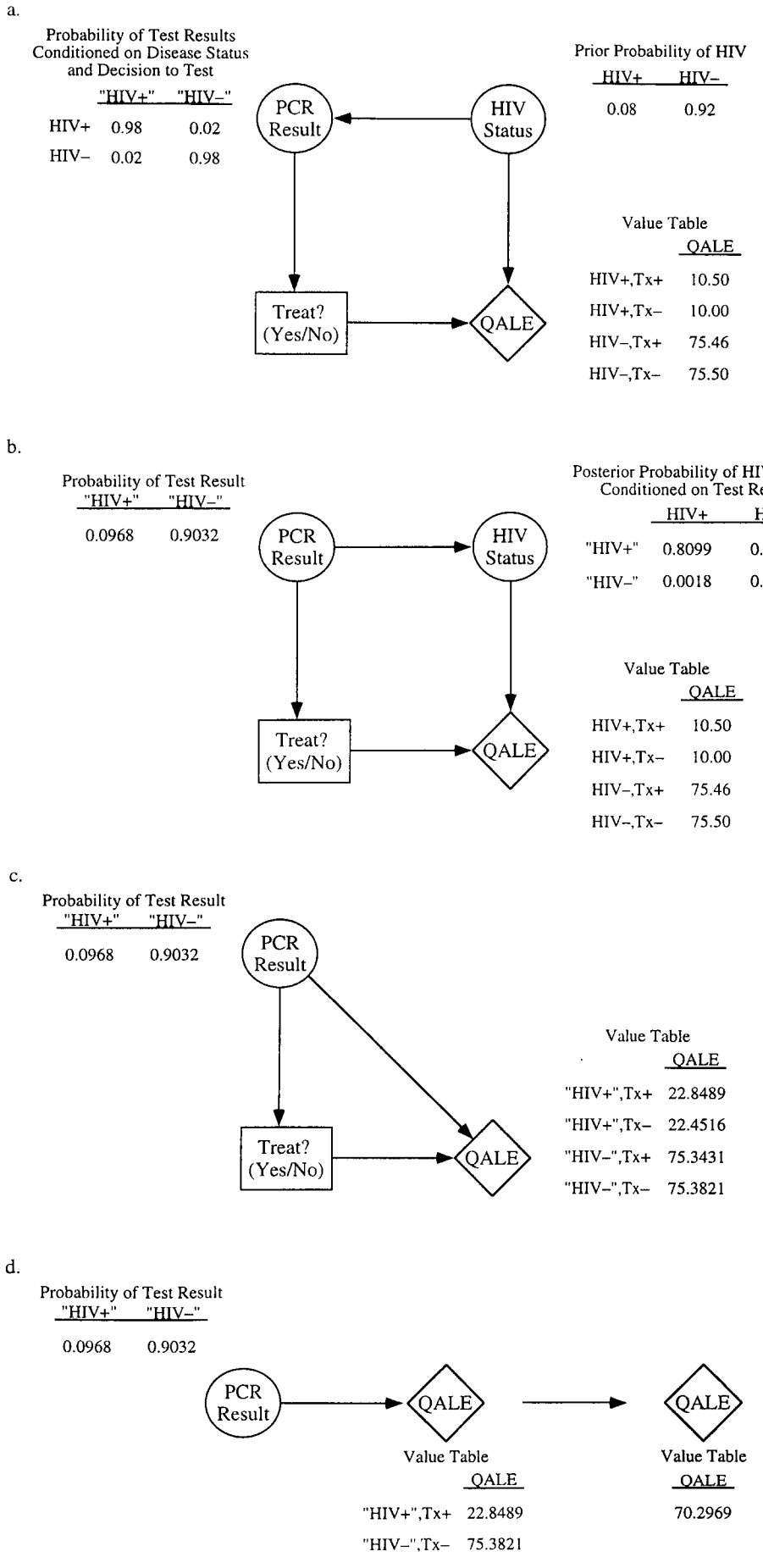


FIGURE 7. The Treat or No Treat Decision with diagnostic test information. HIV = human immunodeficiency virus; HIV+ = HIV-infected; HIV− = not infected with HIV; QALE = quality-adjusted life expectancy; PCR = polymerase chain reaction; Tx+ = treated; Tx− = untreated. Test results are shown in quotation marks ("HIV+") and the true disease state is shown without quotation marks (HIV+).

The initial influence diagram, drawn in assessment order, is shown in (a).

In (b) the influence diagram is drawn in observation order with HIV Status conditioned on PCR Result, as reflected by the direction of the arc between PCR Result and HIV Status, and by the information in the associated probability tables.

HIV Status is removed by averaging (c).

Treat? is removed by policy determination (d). In (d), PCR Result is then removed by averaging.

decision, so we can choose the best policy (more formally, this operation is called maximization; see the appendix). That is, the diagram satisfies the conditions in step 2a, and we can remove the decision node by policy determination (fig. 6c). The Treat alternative has a higher quality-adjusted life expectancy, but the advantage is small. We now have evaluated this influence diagram fully, and have determined that the Treat alternative has the higher expected utility (70.2632 QALYs).

THE TREAT OR NO-TREAT DECISION WITH A DIAGNOSTIC TEST

We now extend the previous example to include diagnostic-test information (for this example, we start with the influence diagram in figure 3b, which we have redrawn as figure 7a). We assume, for this example, that we have already ordered a diagnostic test, and that we shall know the test result at the time we make the Treat? decision. Examining figure 7a, we see that there is no barren node (step 1). We cannot yet make the Treat? decision, because the outcomes of the chance event (HIV Status) are revealed subsequent to the decision (we know that because there is no informational arc from HIV Status to Treat?). Thus, the conditions of step 2a are not satisfied. We cannot remove HIV Status by averaging because it points to PCR Result, thus failing the conditions in step 2b. We notice that, however, if we reverse the arc between HIV Status and PCR Result, the conditions in step 2c are satisfied—we can remove HIV Status by averaging. As shown in figure 7b, we can reverse the arc between PCR Result and HIV Status without adding arcs, because neither node has other parents.

Notice how the information in the probability tables changes when the diagram is changed from assessment order (fig. 7a) to observation order (fig. 7b). The probability table for PCR Result originally contained the sensitivity and specificity of PCR, and the table for HIV Status contained the prior probability of HIV infection (see fig. 7a). In figure 7b, we used Bayes' theorem to reverse the arc and to calculate the probability of a positive or negative test (the denominator of Bayes' theorem provides the probability of a positive or negative test), as shown in the probability table for PCR Result, and the posterior probability of HIV conditioned on positive and negative test result, as shown in the probability table for HIV Status.

After reversing the arc, we return to step 2 of the algorithm. We find that we cannot make the Treat? decision because the outcomes of the chance events represented by HIV Status are revealed subsequent to the Treat? decision; thus, the conditions in step 2a are not satisfied. The conditions of step 2b are satisfied: HIV Status now points to the value node only. Therefore, we can remove HIV Status by averaging (as discussed earlier); that is, we calculate the expected utility for each of the decision alternatives conditioned on the test result. Now, however, HIV Status is conditioned on PCR Result (there is an arc from PCR Result to HIV Status), so when we remove HIV Status, we must draw an arc from PCR Result to QALE, the value node (fig. 7c). Thus, after we remove HIV Status, QALE will still be conditioned on PCR Result and on Treat? We must therefore calculate the entry in the value table for each possible combination of QALE's parents: "HIV+", Tx+; "HIV+", Tx−; "HIV−", Tx+; and "HIV−", Tx−. We use the probabilities from the probability table associated with PCR Result to calculate, for example, the entry in the value table for "HIV+", Tx+ as

$$\begin{aligned}\text{EU}[\text{"HIV+"}, \text{Tx+}] &= p(\text{HIV+} \mid \text{"HIV+"})\text{U}[\text{HIV+}, \text{Tx+}] \\ &\quad + p(\text{HIV−} \mid \text{"HIV+"})\text{U}[\text{HIV−}, \text{Tx+}] \\ &= (0.8099)(10.50\ \text{QALYs}) + (0.1901)(75.46\ \text{QALYs}) \\ &= 22.8489\ \text{QALYs}\end{aligned}$$

Similarly, we can calculate the entry in the value table for "HIV+", Tx− as

$$\begin{aligned}\text{EU}[\text{"HIV+"}, \text{Tx−}] &= p(\text{HIV+} \mid \text{"HIV+"})\text{U}[\text{HIV+}, \text{Tx−}] \\ &\quad + p(\text{HIV−} \mid \text{"HIV+"})\text{U}[\text{HIV−}, \text{Tx−}] \\ &= (0.8099)(10.00\ \text{QALYs}) + (0.1901)(75.50\ \text{QALYs}) \\ &= 22.4516\ \text{QALYs}\end{aligned}$$

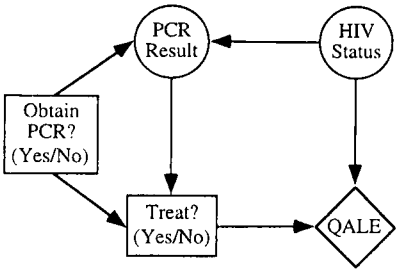
The entries in the value table in figure 7c show the QALEs for the individual strategies.

We return to the algorithm and find that the conditions for step 2a are now satisfied: the decision node Treat? points into the value node, and the only other node that points into the value node (PCR Result) also points into Treat?. Thus, we can remove Treat? by policy determination. Unlike in our previous example, however, at the time that we make the decision, we will have observed a PCR test result. Therefore, we shall continue the example by allowing for either a positive or a negative result to occur, and evaluating the best treatment option conditioned on the test result. We see by inspection of the value table in figure 7c that, as expected, if the PCR result is positive, we obtain a higher quality-adjusted life expectancy by choosing to treat (22.8489 QALYs) than by choosing not to treat (22.4516 QALYs), and if the PCR result is negative, we obtain a higher qual-

a.

Probability of Test Results Conditioned on Disease Status and Decision to Test

		"HIV+"	"HIV−"	"NA"
Obtain PCR	HIV+	0.98	0.02	0
	HIV−	0.02	0.98	0
Do Not Obtain PCR	HIV+	0	0	1.0
	HIV−	0	0	1.0



Prior Probability of HIV

HIV+	HIV−
0.08	0.92

Value Table

	QALE
HIV+,Tx+	10.50
HIV+,Tx−	10.00
HIV−,Tx+	75.46
HIV−,Tx−	75.50

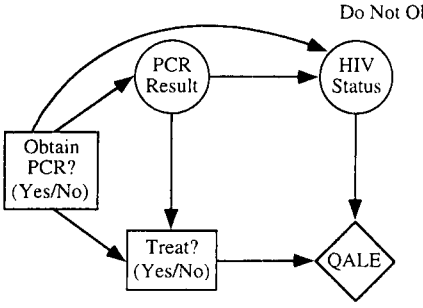
b.

Probability of PCR Result

	"HIV+"	"HIV−"	"NA"
Obtain PCR	0.0968	0.9032	0
Do Not Obtain PCR	0	0	1

Posterior Probability of HIV Status Conditioned on Test Result

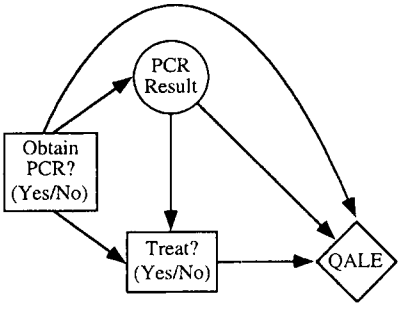
		HIV+	HIV−
Obtain PCR	"HIV+"	0.8099	0.1901
	"HIV−"	0.0018	0.9982
Do Not Obtain PCR	"NA"	0.08	0.92



Value Table

	QALE
HIV+,Tx+	10.50
HIV+,Tx−	10.00
HIV−,Tx+	75.46
HIV−,Tx−	75.50

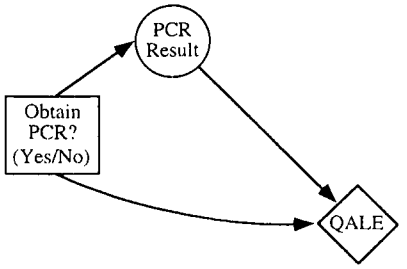
c.



Value Table

		QALE
Obtain PCR	"HIV+",Tx+	22.8489
	"HIV+",Tx−	22.4516
	"HIV−",Tx+	75.3431
	"HIV−",Tx−	75.3821
Do Not Obtain PCR	"NA",Tx+	70.2632
	"NA",Tx−	70.2600

d.



Value Table

		QALE
Obtain PCR	"HIV+"	22.8489
	"HIV−"	75.3821
Do Not Obtain PCR	"NA",Tx+	70.2632

FIGURE 8 *(left and top of facing page).* The Treat, No Treat, or Test decision. HIV = human immunodeficiency virus; HIV+ = HIV-infected; HIV− = not infected with HIV; NA = test result not available; QALE = quality-adjusted life expectancy; PCR = polymerase chain reaction; Tx+ = treated; Tx− = untreated. Test results are shown in quotation marks ("HIV+") and the true disease state is shown without quotation marks (HIV+).

The initial influence diagram, drawn in assessment order, is shown in (a), and depicts the decisions to obtain PCR and then to treat.

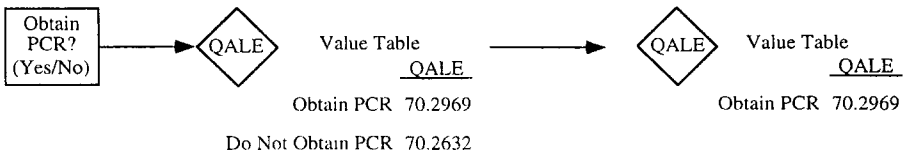
In (b) the diagram is shown in observation order with HIV Status conditioned on PCR Result, as reflected by the direction of the arc between PCR Result and HIV Status, and by the information in the associated probability tables. To reverse the arc, an arc between Obtain PCR and HIV Status has been added.

In (c) HIV Status has been removed by averaging.

In (d) Treat? has been removed by policy determination.

In (e) PCR Result is removed by averaging, then Obtain PCR? is removed by policy determination.

e.



ity-adjusted life expectancy by choosing not to treat. The resulting diagram is shown in figure 7d. Because we continue to show PCR Result as a chance node, the diagram reflects that we have not yet observed the result of the PCR test, which could be either positive or negative. If the test is positive, the expected utility of our decision (treat) is 22.8489 QALYs; if the test is negative, the expected utility of our decision (no treatment) is 75.3821 QALYs. We can calculate the quality-adjusted life expectancy for the average patient who undergoes testing by following the algorithm and removing the chance node by averaging, as shown in figure 7d. We multiply the probability of a positive PCR test result by the expected utility given a positive result, and add this value to the product of the probability of a negative test result and the expected utility given a negative test result, or

$$p(\text{"HIV+"})EU[\text{"HIV+"}, Tx+] + p(\text{"HIV−"})EU[\text{"HIV−"}, Tx−]$$
$$= (0.0968)(22.8489 \text{ QALYs}) + (0.9032)(75.3821)$$
$$= 70.2969 \text{ QALYs}$$

This result indicates that the quality-adjusted life expectancy for an entire cohort of patients (with a prior probability of HIV infection of 0.08) who undergo testing is 70.2969 QALYs.

In summary, to evaluate this diagram, we performed arc reversal using Bayes' theorem, removed the chance node by averaging, and then removed the decision node by policy determination. This sreies of steps left us with a diagram that showed the expected utility conditioned on the PCR test result. In practice, of course, only one result would occur, and we would observe it before making the Treat? decision. We finished the evaluation by removing this final chance node.

THE TREAT OR TEST DECISION

We now return to our opening example, the treat-or-test decision (we have redrawn figure 2 as figure 8a). We begin the evaluation as we did in the example in the previous section. We cannot remove any nodes directly, so we reverse the arc between PCR Result and HIV Status to remove the node by averaging. To reverse this arc, we must add an arc from Obtain PCR? to HIV Status, so that each node has identical parents (fig. 8b). We calculate the probability of possible test results (positive, negative, or not available), and the posterior probability of HIV infection conditioned on each result, as in the previous example, with identical results. If the test is not ordered, the posterior probability for HIV infection is the same as the prior probability, as shown in the probability table for HIV Status. We remove HIV Status (fig. 8c), and, by averaging, calculate the entries in the value table conditioned on PCR Result and Treat?. These results are again identical to those in figure 7c. Referring to the algorithm (step 2a), we notice that we can now remove Treat? by policy determination (fig. 8d). To do so, we recalculate the value table conditioned on the alternatives for Obtain PCR?, and on PCR Result (fig. 8d). Notice that in the value table for figure 8d, we do not show explicitly the preferred policy for each test result (treat [Tx+] if "HIV+", and do not treat [Tx−] if "HIV−"). However, influence-diagram–evaluation software would typically retain information about the correct policy for each test result. Referring to the algorithm again (step 2b), we find that we can now remove PCR Result. We calculate the expected utility for the Obtain PCR alternative by averaging as

$$p(\text{"HIV+"})EU[\text{"HIV+"}, Tx+] + p(\text{"HIV−"})EU[\text{"HIV−"}, Tx−]$$
$$= (0.0968)(22.8489 \text{ QALYs}) + (0.9032)(75.3821)$$
$$= 70.2969 \text{ QALYs}$$

Notice that this calculation is the same as the calculation that we performed at the end of the example in the previous section. The expected utility of the No Test alternative is 70.2632 (calculated in the example in figure 6c), as noted in the value table in figure 8d. We now can evaluate the decision Obtain PCR? and remove the decision node by policy determination (fig. 8e). We see that the expected utility of the Test alternative is higher than that of the No Test (and Treat) alternative, but by only 0.0337 QALYs, which suggests that the decision is a close call. The decision is a close call despite the high sensitivity and specificity of the PCR, primarily because in the absence of the opportunity to test, the optimal alternative is to treat all infants born of HIV-

Table 2 • Assumptions and Data, Inpatient HIV Screening*

Variable	Base Case	Comment
Sensitivity of HIV test	0.995	
Specificity of HIV tests, including confirmatory Western blot	0.999994	Based on rigorously controlled screening programs
Quality-adjusted life expectancy		
HIV−	44.79 years	Life expectancy for a 30-year-old patient
HIV−, treated based on false-positive test result	44.59 years	We assume a loss of 0.2 quality-adjusted years for a false-positive test result followed by unneeded treatment
HIV+, untreated	4.24 years	We assume patients are identified relatively late in the asymptomatic period
HIV+, treated	4.69 years	We assume that treatment provides approximately 1 year of additional life, unadjusted for quality of life. The net difference in quality-adjusted length of life between treated and untreated patients (4.69 − 4.24) depends on both the benefits of early treatment and the effect on quality of life of early detection through screening[15]
Quality-adjusted life gained in sexual partners of patients due to prevention of HIV infection	0.13 year	We assume patients reduce sexual risk behavior by 15% and that each has one partner at risk

*For a detailed list of sources, see Owens et al.[15]

infected mothers, and testing only prevents unnecessary treatment of uninfected infants (which reduces the quality-adjusted life expectancy of the uninfected infants by only 0.04 (QALYs).

SCREENING OF INPATIENTS FOR HIV INFECTION

Recent national guidelines recommend that acute-care hospitals in which the prevalence of HIV infection is 1% or higher offer voluntary HIV screening to all patients.[10] To demonstrate the use of influence diagrams with a problem of realistic complexity, we evaluate the potential effectiveness of such a screening program in a hospital with a prevalence of HIV infection of 5%. We assume that the screening test will be the sequence of antibody tests and confirmatory Western blot typically used in HIV testing in adults, rather than PCR as used in our previous examples (table 2). The influence diagram models two benefits from treatment: a benefit to the person screened (from early medical therapy), and a benefit to the sexual partners (from reduced transmission) of patients who are identified as being infected with HIV. We assume that early medical intervention, including prophylaxis for opportunistic infections and treatment with antiretroviral medication, extends length of life by approximately a year, unadjusted for quality of life.[15,16] We assume that 80% of people identified as having HIV infection reduce their sexual contacts as a result of counseling; we assume these people reduce their sexual contacts by 15%.[15,16] We have estimated previously the net number of HIV infections that would be prevented by such a change in risk behavior from a Markov model that evaluates transmission.[15,16,33] Other assumptions, data, and sources are shown in table 2.

The influence diagram in figure 9a shows the structure of the problem and the relevant data. We use the algorithm to evaluate this diagram, and find that Compliance with Medical Therapy and Reduce Risk Behavior both point into the value node and nowhere else; therefore, they can be removed by averaging. We shall remove Reduce Risk Behavior first, but we could remove either node first. We calculate each entry in the value table by determining that, after we remove Reduce Risk Behavior, QALE is conditioned on HIV Status, on Compliance with Medical Therapy, and on Treat & Counsel? (fig. 9b). Thus, we must calculate an entry in the value table for each combination of the parents for both the Treat option and the No Treat option (HIV+, Comply+; HIV+, Comply−; HIV−, Comply+; HIV−, Comply−), where Comply+ indicates compliance with therapy. For example, we calculate the expected utility for a patient who is HIV+, treated, and complies with therapy as

$$p(\text{Reduce}+)U[\text{HIV}+, \text{Comply}+, \text{Reduce}+] + p(\text{Reduce}-)U[\text{HIV}+, \text{Comply}+, \text{Reduce}-] = (0.8)(4.82 \text{ QALYs}) + (0.2)(4.69 \text{ QALYs}) = 4.794 \text{ QALYs}$$

We remove the node Compliance with Medical Therapy similarly (fig. 9c).

We can evaluate the diagram further using the steps used in the previous section. Although we do not show these steps in figure 9, we note the results

here for readers who wish to complete the example. We reverse the arc between HIV Status and HIV Test Result (by adding an arc from Screen to HIV Status) and calculate $p(\text{"HIV+"}) = 0.0498$, $p(\text{"HIV−"}) = 0.9502$, $p(\text{HIV+} \mid \text{"HIV+"}) = 0.9999$, $p(\text{HIV+} \mid \text{"HIV−"}) = 0.0003$. The evaluation shows that, in the absence of testing, not treating provides a higher quality-adjusted life expectancy (42.7625 QALYs) than does treating (42.6170 QALYs), and that screening results in a small but positive gain in quality-adjusted life expectancy (42.7846 QALYs) relative to not screening and not treating (42.7625 QALYs). Although the gain in QALYs per patient is small in our simplified example, even with this gain, if one million patients

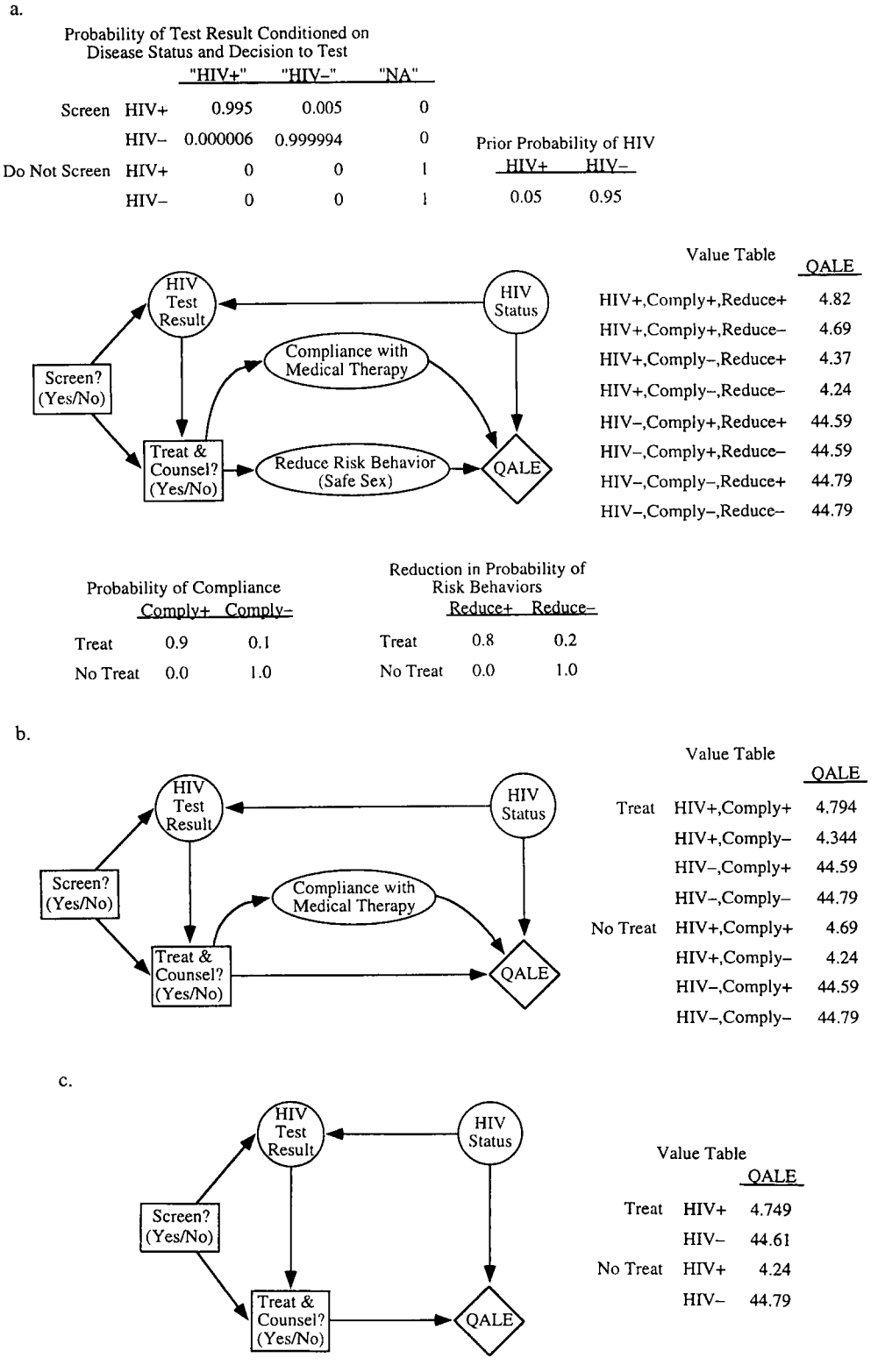


FIGURE 9. Screening of inpatients for HIV infection. HIV = human immunodeficiency virus; HIV+ = HIV-infected; HIV− = not infected with HIV; QALE = quality-adjusted life expectancy; Tx+ = treated; Tx− = untreated. Test results are shown in quotation marks ("HIV+") and the true disease state is shown without quotation marks (HIV+).

(a) The influence diagram for the decision to screen and then treat inpatients for HIV infection, drawn in assessment order.

In (b) Reduce Risk Behavior has been removed by averaging. Because influence diagrams assume structural symmetry, the value table has entries for two logically impossible combinations (No Treat, HIV+, Comply+ and No Treat, HIV−, Comply+). These values are not used as the influence diagram is further evaluated.

In (c) Compliance with Medical Therapy has been removed by averaging.

were screened, such a program could result in a gain of over 20,000 life years (see Owens et al.[15] for a more detailed analysis).

EXTENSIONS TO THE EVALUATION PROCESS

Our examples show how to evaluate the influence diagram with a specific set of values for each variable. We can use the influence diagram to perform sensitivity analyses as well. The method is identical to that used with decision trees. We vary the value of a variable over a plausible range, and evaluate the influence diagram for each specific value of the variable. For example, to perform a two-way sensitivity analysis on Compliance with Medical Therapy and Reduce Risk Behavior (fig. 9), we evaluate the diagram after changing the value of each variable, and repeat this process iteratively until we have examined each of the plausible combinations of values for these variables. The evaluation algorithms can perform the calculations for the sensitivity analysis efficiently by reducing the diagram without removing the node on which the sensitivity analysis is being performed, thus reducing the required computations.

We also note an additional extension to the evaluation algorithm described above. That algorithm will evaluate influence diagrams with deterministic nodes correctly, but does not exploit the special properties of deterministic nodes to reduce computational requirements. More sophisticated algorithms can exploit the special properties of deterministic nodes to achieve computational efficiencies.

Strengths and Limitations of Influence Diagrams

Influence diagrams represent the relationships among variables. These relationships are important because they reflect the analyst's, or the content expert's, knowledge about a problem. For example, in figure 4, the nodes and arcs (and absence of arcs) in the top half of the diagram represent the variables and probabilistic relationships that mediate the health benefit of an HIV screening program. The construction of such a model often involves a collaboration between an analyst and a content expert. This collaboration represents an exercise in knowledge acquisition—the analyst attempts to construct a model that reflects the content expert's understanding of the medical domain. In building the diagram of figure 4, the analyst might first ask a content expert for the major variables that affect health outcomes (Surgical Transmission, Transmission to Sexual Partners, and Early-Treatment Benefit). The analyst could then discuss the probabilistic relationships among the variables, and the factors that, in turn, affect each of the identified variables. By building the graphic elements of the influence diagram, the analysts can focus on the relationships among variables before adding the detail needed for the associated probability and value tables. For problems in which the probabilistic relationships are complex, or have special importance, influence diagrams may be a useful aid in this knowledge-acquisition task.

A second feature of influence diagrams, as mentioned earlier, is that they can be drawn with conditioning displayed in the manner that most facilitates assessment of the probabilities. This feature may also facilitate knowledge acquisition. For example, a clinician expert may be able to assess the prevalence of disease and the sensitivity and specificity of a diagnostic test more easily than he or she could assess the posttest probability of disease. After the influence diagram is drawn to facilitate probability assessments, all updating and Bayesian inference are handled automatically by the evaluation algorithms. Although there are approaches for performing Bayesian updating within a decision tree, for problems with extensive Bayesian updating, such as sequential-testing decisions, influence diagrams ease the burden on the analyst by reducing the need for complex equations required for Bayesian updating in the tree. Influence diagrams also reduce the time required to find errors that may be introduced when these equations are specified.

Although influence diagrams offer advantages for certain analytic problems, they also have limitations relative to the decision-tree format. Highly asymmetric problems may be easier to understand when represented as decision trees, as mentioned earlier. The timing of events may be easier to identify in a decision tree, although the same information is explicit in the influence diagram. In addition, it is not enough to simply draw the influence diagram. To represent fully the decision alternatives, strategies, alternative events, and values of outcomes, the analyst must complete the probability and value tables in the influence diagram, a process similar to placing the values of variables in the decision tree.

As with a decision tree, an influence diagram for a complex problem may require a large number of probability assessments. For an asymmetric decision problem, the analyst must use probabilities of 0 or 1 to represent the asymmetry. Although there are methods for reducing the number of probability assessments needed to specify an influence diagram,[34] the size and complexity of the probability and value tables also increase rapidly as the problem modeled becomes more complex (see, for example, the value table in figure 9). The probability tables also become substantially larger if the chance events have multiple outcomes (for simplicity, we

used dichotomous outcomes in most of our examples, but chance nodes may have as many possible outcomes as the analyst chooses). In general, if chance node A has n outcomes, and k parents, and ith parent has m_i outcomes, the number of entries in the probability table is

$$n \prod_{i=1}^{k} m_i$$

For example, if chance node A has three outcomes ($n = 3$), and three parents ($k = 3$), each with three outcomes ($m_i = 3$, for $i = 1, 2, 3$), then the number of entries in the probability table is $3 \cdot 3 \cdot 3 \cdot 3 = 81$.

In addition, the evaluation algorithms for influence diagrams are designed for computer-based implementation. A moderate-sized tree can be solved by hand; only the simplest of influence diagrams could be solved readily without software. As arcs are reversed during evaluation of complex diagrams, extra arcs are added and the diagram may become confusing. Although software programs that evaluate influence diagrams have been lacking, several programs now make use of influence diagrams. However, representation of dynamic models (for example, Markov models) is possible[35,36] but is not implemented in currently available software. Finally, certain problems can be represented and evaluated more easily with closed-form algebraic solutions[37] or with more general functional relationships than with influence diagrams; other problems can be solved by direct manipulation of these tables.

Conclusions

Influence diagrams are a powerful graphic representation for decision models, complementary to decision trees. They are particularly helpful when problems have a high degree of conditional independence, when compact representation of extremely large models is needed, when communication of the probabilistic relationshps is important, or when the analysis requires extensive Bayesian updating. Influence diagrams provide a useful alternative tool for the analyst. The choice of graphic representation should be governed by convenience, and will depend on the problem being analyzed, on the experience of the analyst, and on the background of the consumers of the analysis.

The authors thank Stephen Pauker, Lyn Dupré, Blackford Middleton, John Hornberger, Gillian Sanders, Mark Helfand, and Anne Dembitzer for helpful comments about the manuscript; and Andria Cardinalli for preparing the figures.

References

1. Miller AC, Merkofer MM, Howard RA, Matheson JE, Rice TR. Development of automated aids for decision analysis. Menlo Park, CA: Stanford Research Institute, 1976.
2. Shachter RD, Eddy DM, Hasselblad V. An influence diagram approach to the confidence profile method for health technology assessment. Proceedings of the Conference on Influence Diagrams for Decision Analysis, Inference and Prediction. University of California, Berkeley, CA, 1988:299–306.
3. Eddy DM, Hasselblad V, Shachter R. Meta-Analysis by the Confidence Profile Method: The Statistical Synthesis of Evidence. Boston, MA: Academic Press, 1992.
4. Lehmann HP, Shortliffe EH. THOMAS: Building Bayesian statistical expert systems to aid in clinical decision making. In: Miller RA (ed). Proceedings of the Symposium on Computer Applications in Medical Care. Los Alamitos, CA: IEEE Computer Society Press, 1990:58–64.
5. Lehmann HP. Computational formulation of Bayesian statistical models for interpreting clinical research literature [PhD thesis]. Section on Medical Informatics, Department of Medicine, Stanford University, Stanford, CA, 1991.
6. Jimison HB. A representation for gaining insight into clinical decision models [PhD thesis]. Section on Medical Informatics, Department of Medicine, Stanford University, Stanford, CA, 1990.
7. Owens DK, Nease RF. Development of outcome-based practice guidelines: a method for structuring problems and synthesizing evidence. Joint Commission Journal on Quality Improvement. 1993;19:248–63.
8. Suermondt HJ. Explanation in Bayesian belief networks [PhD thesis]. Section on Medical Informatics, Department of Medicine, Stanford University, Stanford, CA, 1991.
9. Nease RF Jr, Owens DK. Use of influence diagrams to structure medical decisions. Med Decis Making. 1997;17:263–75.
10. Centers for Disease Control. 1995 revised guidelines for prophylaxis against *Pneumocystis carinii* pneumonia for children infected with or perinatally exposed to human immunodeficiency virus. MMWR. 1995;44:1–11.
11. Charniak E. Bayesian networks without tears. AI Magazine. 1991;Winter:50–63.
12. Shwe MA, Middleton B, Heckerman DE, et al. Probabilistic diagnosis using a reformulation of the INTERNIST-1/QMR knowledge base. I. The probabilistic model and inference algorithms. Methods Inf Med. 1991;30:241–55.
13. Middleton B, Shwe MA, Heckerman DE, et al. Probabilistic diagnosis using a reformulation of the INTERNIST-1/QMR knowledge base. II. Evaluation of diagnostic performance. Methods Inf Med. 1991;30:256–7.
14. Owens DK, Holodniy M, McDonald TW, Scott J, Sonnad S. A meta-analytic evaluation of the polymerase chain reaction for the diagnosis of HIV infection in infants. JAMA. 1996;275: 1342–8.
15. Owens DK, Nease RF, Harris RA. Cost–effectiveness of HIV screening in acute care settings. Arch Intern Med. 1996;156: 394–404.
16. Owens DK, Harris RA, Scott PM, Nease RF. Screening surgeons for human immunodeficiency virus (HIV). A cost–effectiveness analysis. Ann Intern Med. 1995;122:641–52.
17. Raiffa H. Decision Analysis. Reading, MA: Addison–Wesley, 1968.
18. Shachter RD. An ordered examination of influence diagrams. Networks. 1990;20:535–63.
19. Howard RA, Matheson JE. Influence diagrams. In: Howard RA, Matheson JE (eds). The Principles and Applications of Decision Analysis. Vol. II. Menlo Park, CA: Strategic Decisions Group, 1984:720–62.
20. Shachter RD. Probabilistic inference and influence diagrams. Oper Res. 1988;36:589–605.

21. Owens DK, Holodniy M, Garber AM, et al. Polymerase chain reaction for the diagnosis of HIV infection in adults. A meta-analysis with recommendations for clinical practice and study design. Ann Intern Med. 1996;124:803–15.
22. Shachter RD. Evaluating influence diagrams. Oper Res. 1986; 34:871–882.
23. Olmsted SM. On representing and solving decision problems [PhD thesis]. Department of Engineering–Economic Systems, Stanford University, Stanford, CA, 1983.
24. Jensen FV, Olesen KG, Andersen SK. An algebra of Bayesian belief universes for knowledge based systems. Networks. 1990;20:637–59.
25. Jensen FV, Lauritzen SL, Olesen KG. Bayesian updating in causal probabilistic networks by local computations. Comp Stat Q. 1990;4:269–82.
26. Andersen SK, Olesen KG, Jensen FV, Jensen F. HUGIN: A shell for building belief universes for expert systems. Proceedings of the Eleventh International Joint Conference on Artificial Intelligence. Detroit, MI, 1989:1080–5.
27. Lauritzen SL, Spiegelhalter DJ. Local computations with probabilities on graphical structures and their application to expert systems. J R Stat Soc. 1988;50:157–224.
28. Shafer G, Shenoy PP. Probability propagation. Ann Math Artificial Intel. 1990;2:327–52.
29. Shachter RD, Andersen SK, Poh KL. Directed reduction algorithms and decomposable graphs. In: Bonnisone P, Henrion M, Kanal LN, Lemmer JF (eds). Uncertainty in Artificial Intelligence 6. Amsterdam, The Netherlands: North Holland, 1991:197–208.
30. Ndilikilikesha P. Potential influence diagrams. [Working Paper 235]: School of Business, University of Kansas, Lawrence, KS, 1991.
31. Shachter RD, Peot MA. Decision making using probabilistic inference methods. In: Uncertainty in Artificial Intelligence: Proceedings of the Eighth Conference. San Mateo, CA: Morgan Kaufmann, 1992:276–83.
32. Shenoy PP. A fusion algorithm for solving Bayesian decision problems. In: D'Ambrosio B, Smets P, Bonissone P (eds). Uncertainty in Artificial Intelligence: Proceedings of the Seventh Conference. San Mateo, CA: Morgan Kaufmann, 1991:361–9.
33. Owens DK, Nease RF. Transmission of human immunodeficiency virus (HIV) infection between physicians and patients: a model-based analysis of risk. In: Kaplan E, Brandeau ML (eds). Modeling the AIDS Epidemic: Planning, Policy and Prediction. New York: Raven Press, 1994:153–77.
34. Heckerman DE. Probabilistic similarity networks [PhD thesis]. Section on Medical Informatics, Department of Medicine, Stanford University, Stanford, CA, 1990.
35. Tatman JA. Decision processes in influence diagrams: formulation and analysis [PhD thesis]. Department of Engineering–Economic Systems, Stanford University, Stanford, CA, 1985.
36. Tatman JA, Shachter RD. Dynamic programming and influence diagrams. IEEE SMC. 1990;20:365–79.
37. Pauker SG, Kassirer JP. The threshold approach to clinical decision making. N Engl J Med. 1980;302:1109–17.

APPENDIX

This appendix further explains the mathematical operations that allow inference and decision making in both decision trees and influence diagrams. These operations form the basis of the solution algorithms for influence diagrams. Because the joint probability distribution plays a central role in the probabilistic operations, we begin with it.

Joint probability distributions. Figure A1 shows a joint probability distribution for the events "PCR Result" and "HIV Status," calculated from the conditional and prior probabilities shown in figure 2. Notice that an entry in the table denotes the probability that a particular test result and an infection status *both* occur. Unlike in the conditional probability tables presented in figures 1 through 9, here the sum of the probabilities in all the cells of the joint probability distribution is 1, and the sum of a row or column provides the unconditioned, or marginal, probability of that event. For example, the probability that the test result is positive *and* the patient is infected is 0.0784 (as distinct from the probability that the test result is positive *given* that the person is infected, which is shown to be 0.98 in figure 2); the probability of a positive test result is 0.0968, which we obtain by summation across the top row of the table.

To understand the evaluation algorithms, we must understand the basic probabilistic operations on the joint probability distribution that underlie the decision-analytic model. Suppose that all the variables in a problem are partitioned into two sets, A and B, and that the utility is given by U. Then, the different operations are

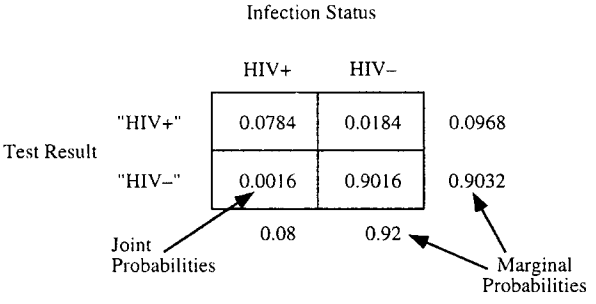


FIGURE A1. The joint probability distribution. The joint probabilities are denoted by the cell entries. The marginal probabilities are obtained by summation of a row or column. HIV = human immunodeficiency virus; HIV+ = HIV-infected; HIV− = not infected with HIV. Test results are shown in quotation marks ("HIV+") and the true disease state is shown without quotation marks (HIV+).

Factorization	$p(AB) = p(A)p(B\|A)$
Marginalization	$p(A) = \Sigma_b\, p(AB)$
Conditionalization	$p(B\|A) = p(AB)/p(A)$
Expectation	$EU\|A = \Sigma_b\, (EU\|AB)p(B\|A)$
Maximization	optimal $EU\|A = \max_b EU\|AB$, (where B is chosen by the decision maker after he or she observes A)

In marginalization, conditionalization, and expectation, *b* represents the possible outcomes for the chance events. These basic operations are sufficient to derive all the de-

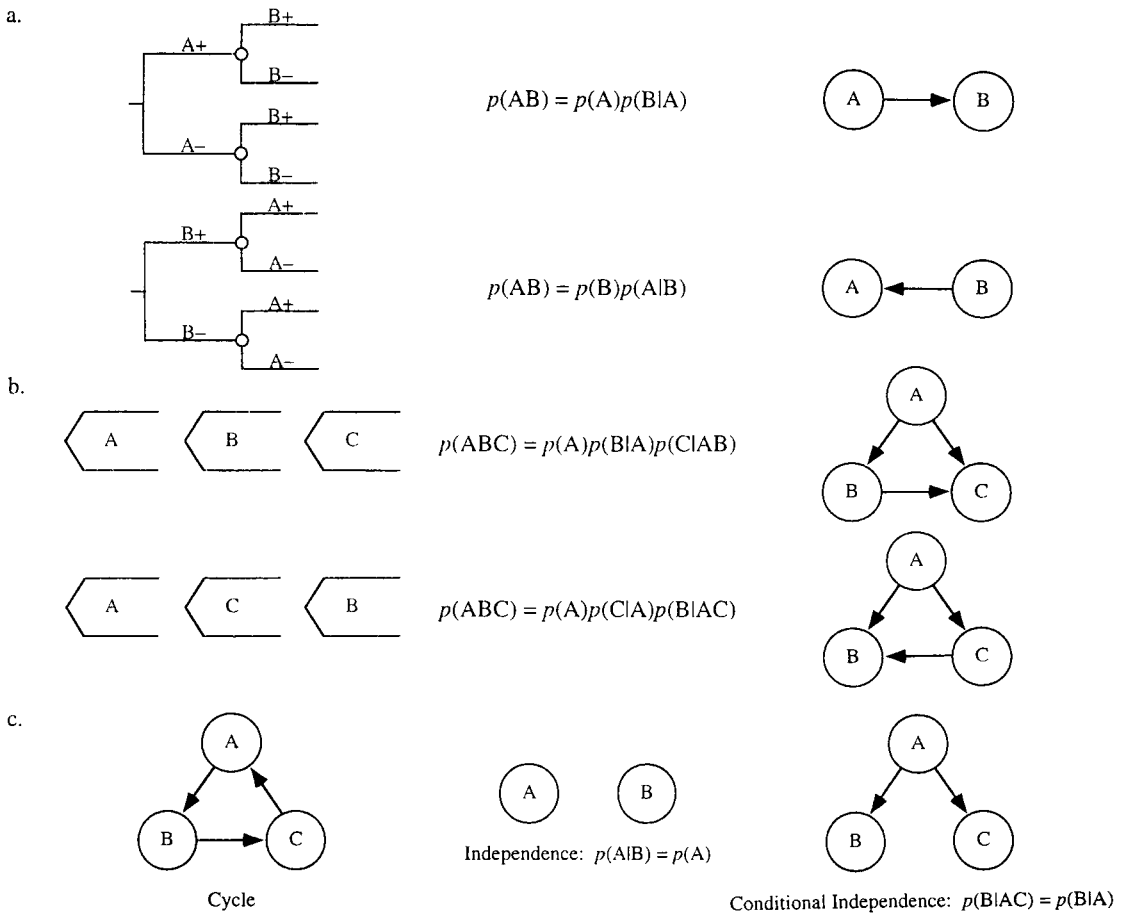


FIGURE A2. Graphic representations of probabilistic conditioning. Panels (a) and (b) are graphic representations of different factorizations of the joint probability space. In the top tree in (a), B is conditioned on A; in the bottom tree, A is conditioned on B. A change in the order of conditioning in the influence diagram is indicated by a change in the direction of the arc. Panel (c) shows a cycle, and how independence and conditional independence are represented in the influence diagram. In the right diagram, nodes B and C are both connected to A, but are not connected directly to each other. This diagram indicates that events B and C are conditionally independent, given that A is known. [Event B is conditionally independent of C, given A, if, when A is known, learning about C provides no information about B, $p(B|AC) = p(B|A)$.] In the influence diagram, this independence is represented by the absence of an arc connecting B and C.

cision-tree and influence-diagram operations used in this paper; we can use them to derive any of the other solution algorithms.[29] Performing chance-node removal by averaging is simply taking the expectation. Decision-node removal by policy determination is maximization. Notice also that the first three operations are the basis of Bayes' theorem. When variables are in order of observation (as they are in a decision tree), no further processing with Bayes' theorem is needed. Therefore, only the final two operations (expectation and maximization) are required for evaluation of a decision tree.

The key to probabilistic inference is that the joint probability distribution can be factored (or expanded) in equivalent but different ways (for an extended example and discussion of these concepts, see Howard and Matheson[19]). From the definition of conditional probability, $p(A|B) = p(AB)/p(B)$, and from the fact that $p(AB) = p(BA)$, we have

$$p(AB) = p(A)p(B|A) = p(B)p(A|B) \quad \text{(A1)}$$

or, for three events,

$$\begin{aligned} p(ABC) &= p(A)p(B|A)p(C|AB) = p(A)p(C|A)p(B|AC) \\ &= p(C)p(A|C)p(B|AC) = p(C)p(B|C)p(A|BC) \\ &= p(B)p(A|B)p(C|AB) = p(B)p\ (C|B)p(A|BC) \end{aligned} \quad \text{(A2)}$$

The operation of moving between expansions is performed with Bayes' theorem. Figures A2a and A2b show examples of the expansions of the joint probability distribution given by equations A1 and A2, and the corresponding influence diagrams and decision trees. Using different expansions corresponds to reordering the events in the decision tree or to changing the direction of the arcs in the influence diagram. For example, at the top of figure A2a, B is conditioned on A. Suppose that A corresponds to the disease state, and that B corresponds to a test result. Then, $p(B+|A+)$ is the sensitivity of the test, and $p(B-|A-)$ is the specificity of the test. Notice that this tree, and the corresponding influence diagram, are drawn in

assessment order. The tree could also be redrawn in *order of observation*, as shown in the lower tree in figure A2a; the necessary conditional probabilities are calculated with Bayes' theorem.

Figure A2c shows an expansion of the joint distribution, called a cycle, that is not allowed. A cycle occurs in an influence diagram when there is a path along the arcs such that we can start from node A and return to it by following the arcs. A cycle is not allowed in an influence diagram because it does not allow an appropriate expansion ordering of the variables. The cycle shows that A is conditioned on C, which is conditioned on B, which is conditioned on A. Such an ordering implies that $p(ABC) = p(A|B)p(B|A)p(C|A)$, which violates the definition of the joint probability. Figure A2c also shows the influence-diagram representation of probabilistic independence, and conditional independence. Conditional independence between B and C is indicated by the absence of an arc.

We now illustrate how we can use these mathematical operations to solve the problem illustrated in figure 7. The first step in the evaluation was to reverse the arc with Bayes' theorem (figure 7a and figure 7b). Thus, our objective is to calculate the probability of disease conditioned on the test results. To do this calculation, we must use factorization to find the joint probability distribution, marginalization to find the marginal probabilities, and then conditionalization to calculate the desired conditional probabilities. We begin by multiplying the conditional probabilities and prior probabilities (noted in figure 7a) to obtain the joint probability distribution (factorization):

$$\begin{aligned} p(\text{"HIV+"}, \text{HIV+}) &= p(\text{HIV+})p(\text{"HIV+"}|\text{HIV+}) \\ &= (0.08)(0.98) = 0.0784 \\ p(\text{"HIV+"}, \text{HIV−}) &= p(\text{HIV−})p(\text{"HIV+"}|\text{HIV−}) \\ &= (0.92)(0.02) = 0.0184 \\ &= p(\text{"HIV+"})p(\text{HIV+}|\text{"HIV+"}) \end{aligned}$$

We calculate the other joint probabilities similarly (fig. A1). Next, we calculate the marginal probabilities (marginalization) from the joint probabilities in the cells of the table in figure A1:

$$\begin{aligned} p(\text{"HIV+"}) = \Sigma_{\text{HIV}} &= p(\text{"HIV+"}, \text{HIV+}) + p(\text{"HIV+"}, \text{HIV−}) \\ &= 0.0784 + 0.0184 = 0.0968 \end{aligned}$$

Finally, we conditionalize:

$$\begin{aligned} p(\text{HIV+}|\text{"HIV+"}) &= p(\text{"HIV+"}, \text{HIV+})/p(\text{"HIV+"}) \\ &= (0.0784)/(0.0968) = 0.8099 \end{aligned}$$

With these three operations (which are the basis of Bayes' theorem), we can calculate any probability desired, such as the posttest probability of disease or the probabilities of a positive or of a negative test result.

The final two operations needed for the influence-diagram–solution algorithms are expectation and maximization. Expectation and maximization are familiar from the evaluation of decision trees. More generally, we find the expected utility, given an event A+, as follows:

$$\text{EU}|\text{A+} = \Sigma_b\, (\text{EU}|\text{A+}, \text{B})p(\text{B}|\text{A+})$$

Thus, in the decision tree in figure A2a, the expected utility at A+ is simply

$$\text{EU}|\text{A+} = \text{U[A+, B+]}p(\text{B+}|\text{A+}) + \text{U[A+, B−]}p(\text{B−}|\text{A+})$$

The final operation, maximization, is denoted as

$$\text{Optimal EU}|\text{A+} = \max_b \text{EU}|\text{A+, B}$$

In this equation, A+ represents an event (for example, a test result) that has been observed by the decision maker before the decision, and B represents the decision alternatives (for example, treatment alternatives). The equation indicates that the optimal choice is the alternative that has the highest expected utility. Thus, if A+ has been observed, the decision maker chooses the alternative (B+ or B−) with the highest utility {U[A+, B+], or U[A+, B−]}.