MARKOV MODELS IN MEDICAL DECISION MAKING: A PRACTICAL GUIDE

FRANK A. SONNENBERG, MD
J. ROBERT BECK, MD

Reprinted from MEDICAL DECISION MAKING, ***Published by***
HANLEY & BELFUS, Inc., Philadelphia, PA

Markov Models in Medical Decision Making:

A Practical Guide

FRANK A. SONNENBERG, MD, J. ROBERT BECK, MD

Markov models are useful when a decision problem involves risk that is continuous over time, when the timing of events is important, and when important events may happen more than once. Representing such clinical settings with conventional decision trees is difficult and may require unrealistic simplifying assumptions. Markov models assume that a patient is always in one of a finite number of discrete health states, called Markov states. All events are represented as transitions from one state to another. A Markov model may be evaluated by matrix algebra, as a cohort simulation, or as a Monte Carlo simulation. A newer representation of Markov models, the Markov-cycle tree, uses a tree representation of clinical events and may be evaluated either as a cohort simulation or as a Monte Carlo simulation. The ability of the Markov model to represent repetitive events and the time dependence of both probabilities and utilities allows for more accurate representation of clinical settings that involve these issues. ***Key words:*** Markov models; Markov-cycle decision tree; decision making. (Med **Decis Making 1993;13:322-339)**

A decision tree models the prognosis of a patient subsequent to the choice of a management strategy. For example, a strategy involving surgery may model the events of surgical death, surgical complications, and various outcomes of the surgical treatment itself. For practical reasons, the analysis must be restricted to a finite time frame, often referred to as the ***time horizon*** of the analysis. This means that, aside from death, the outcomes chosen to be represented by terminal nodes of the tree may not be final outcomes, but may simply represent convenient stopping points for the scope of the analysis. Thus, every tree contains terminal nodes that represent "subsequent prognosis" for a particular combination of patient characteristics and events.

There are various ways in which a decision analyst can assign values to these terminal nodes of the decision tree. In some cases the outcome measure is a crude life expectancy; in others it is a ***quality-adjusted*** life expectancy.' One method for estimating life expectancy is the declining exponential approximation of life expectancy (DEALE),[2] which calculates a patient-specific mortality rate for a given combination of patient characteristics and comorbid diseases. Life expectancies may also be obtained from Gompertz models of survival[3] or from standard life tables.[4] This paper explores another method for estimating life expectancy, the Markov model.

In 1983, Beck and Pauker described the use of Markov models for determining prognosis in medical applications? Since that introduction, Markov models have been applied with increasing frequency in published decision analyses.[6-9] Microcomputer software has been developed to permit constructing and evaluating Markov models more easily. For these reasons, a revisit of the Markov model is timely. This paper serves both as a review of the theory behind the Markov model of prognosis and as a practical guide for the construction of Markov models using microcomputer decision-analytic software.

Markov models are particularly useful when a decision problem involves a risk that is ongoing over time. Some clinical examples are the risk of hemorrhage while on anticoagulant therapy, the risk of rupture of an abdominal aortic aneurysm, and the risk of mortality in any person, whether sick or healthy. There are two important consequences of events that have ongoing risk. First, the times at which the events will occur are uncertain. This has important implications because the utility of an outcome often depends on when it occurs. For example, a stroke that occurs immediately may have a different impact on the patient than one that occurs ten years later. For economic analyses, both costs and utilities are discounted[10,11] such that later events have less impact than earlier ones. The second consequence is that a given event may occur more than once. As the following example shows, representing events that are repetitive or that occur with uncertain timing is difficult using a simple tree model.

Received February 23,1993, from the Division of General Internal Medicine, Department of Medicine, UMDNJ Robert Wood Johnson Medical School, New Brunswick, New Jersey (FAS) and the Information Technology Program, Baylor College of Medicine, Houston, Texas (JRB). Supported in part by Grant LMO5266 from the National Library of Medicine and Grant HS06396 from the Agency for Health Care Policy and Research.

Address correspondence and reprint requests to Dr. Sonnenberg: Division of General Internal Medicine, UMDNJ Robert Wood Johnson Medical School, 97 Paterson Street, New Brunswick, NJ 08903.

A Specific Example

Consider a patient who has a prosthetic heart valve and is receiving anticoagulant therapy. Such a patient may have an embolic or hemorrhagic event at any time. Either kind of event causes morbidity (short-term and/or chronic) and may result in the patient's death. The decision tree fragment in figure 1 shows one way of representing the prognosis for such a patient. The first chance node, labelled **ANTICOAG**, has three branches, labelled BLEED, EMBOLUS, and **NO EVENT.** Both **BLEED** and **EMBOLUS** maybe either **FATAL** or **NON-FATAL**. If NO **EVENT occurs,** the patient remains **WELL**.

There am several shortcomings with this model. First, the model does not specify when events occur. Second, the structure implies that either hemorrhage or embolus may occur only once. In fact, either may occur more than once. Finally, at the terminal nodes labelled **POSTEMBOLUS, POSTBLEED,** and **WELL,** the analyst still is faced with the problem of assigning utilities, a task equivalent to specifying the prognosis for each of these non-fatal outcomes.

The first problem, specifying when events occur, may be addressed by using the tree structure in figure 1 and making the assumption that either **BLEED**or **EMBOLUS** occurs at the average time consistent with the known rate of each complication. For example, if the rate of hemorrhage is a constant 0.05 per person **per year,** then the average time before the occurrence of a hemorrhage is 1/0.05 or 20 years. Thus, the event of having a fatal hemorrhage will be associated with a utility of 20years of normal-quality survival. However, the patient's normal life expectancy may be less than 20 years. Thus, the occurrence of a stroke would have the paradoxical effect of improving the patient's life expectancy. Other approaches, such as assuming that the stroke occurs halfway through the patient's normal life expectancy, are arbitrary and may lessen the fidelity of the analysis.

Both the timing of events and the representation of events that may occur more than once can be addressed by using a recursive decision tree.[12] In a recursive tree, some nodes have branches that have appeared previously in the tree. Each repetition of the tree structure represents a convenient length of time and any event may be considered repeatedly. A recursive tree that models the anticoagulation problem is depicted in figure 2.

Here, the nodes representing the previous terminal nodes **POST-BLEED, POST-EMBOLUS,** and **NO EVENT** are replaced by the chance node **ANTICOAG,** which appeared previously at the root of the tree. Each occurrence of **BLEED** or **EMBOLUS** represents a distinct time period, so the recursive model can represent when events occur. However, despite this relatively simple model and carrying out the recursion for only two time periods, the tree in figure 2 is "bushy," with 17 terminal branches. If each level of recursion represents one year, then carrying out this analysis for even five years would result in a tree with hundreds of terminal branches. Thus, a recursive model is tractable only for a very short time horizon.

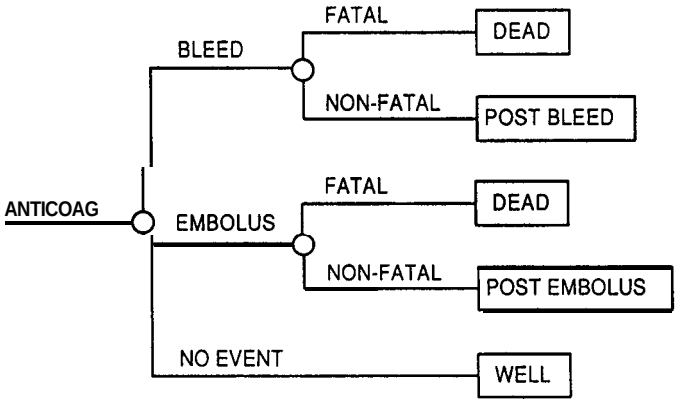


FIGURE 1. Simple tree fragment modeling complications of anticoagulant therapy.

The Markov Model

The Markov model provides a far more convenient way of modelling prognosis for clinical problems with ongoing risk. The model assumes that the patient is always in one of a finite number of states of health referred to as ***Markov states.*** All events of interest are modelled as transitions from one state to another. Each state is assigned a utility, and the contribution of this utility to the overall prognosis depends on the length of time spent in the state. In our example of a patient with a prosthetic heart valve, these states are **WELL, DISABLED,** and **DEAD**. For the sake of simplicity in this example, we assume that either a bleed or a non-fatal embolus will result in the same state **(DISABLED)** and that the disability is permanent.

The time horizon of the analysis is divided into equal increments of time, referred to as Markov cycles. During each cycle, the patient may make a transition from one state to another. Figure 3 shows a commonly used representation of Markov processes, called a ***state-transition diagram,*** in which each state is represented by a circle. Arrows connecting two different states indicate allowed transitions. Arrows leading from a state to itself indicate that the patient may remain in that state in consecutive cycles. Only certain transitions are allowed. For example, a person in the **WELL** state transition from **DISABLED** to **WELL** is not allowed. A person in either the **WELL** state or the **DISABLED** state may die and thus make a transition to the **DEAD** state. However, a person who is in the **DEAD** state, obviously, cannot make a transition to any other state. Therefore, a single armw emanates from the **DEAD** state, leading back to itself. It is assumed that a patient in a given state can make only a single state transition during a cycle.

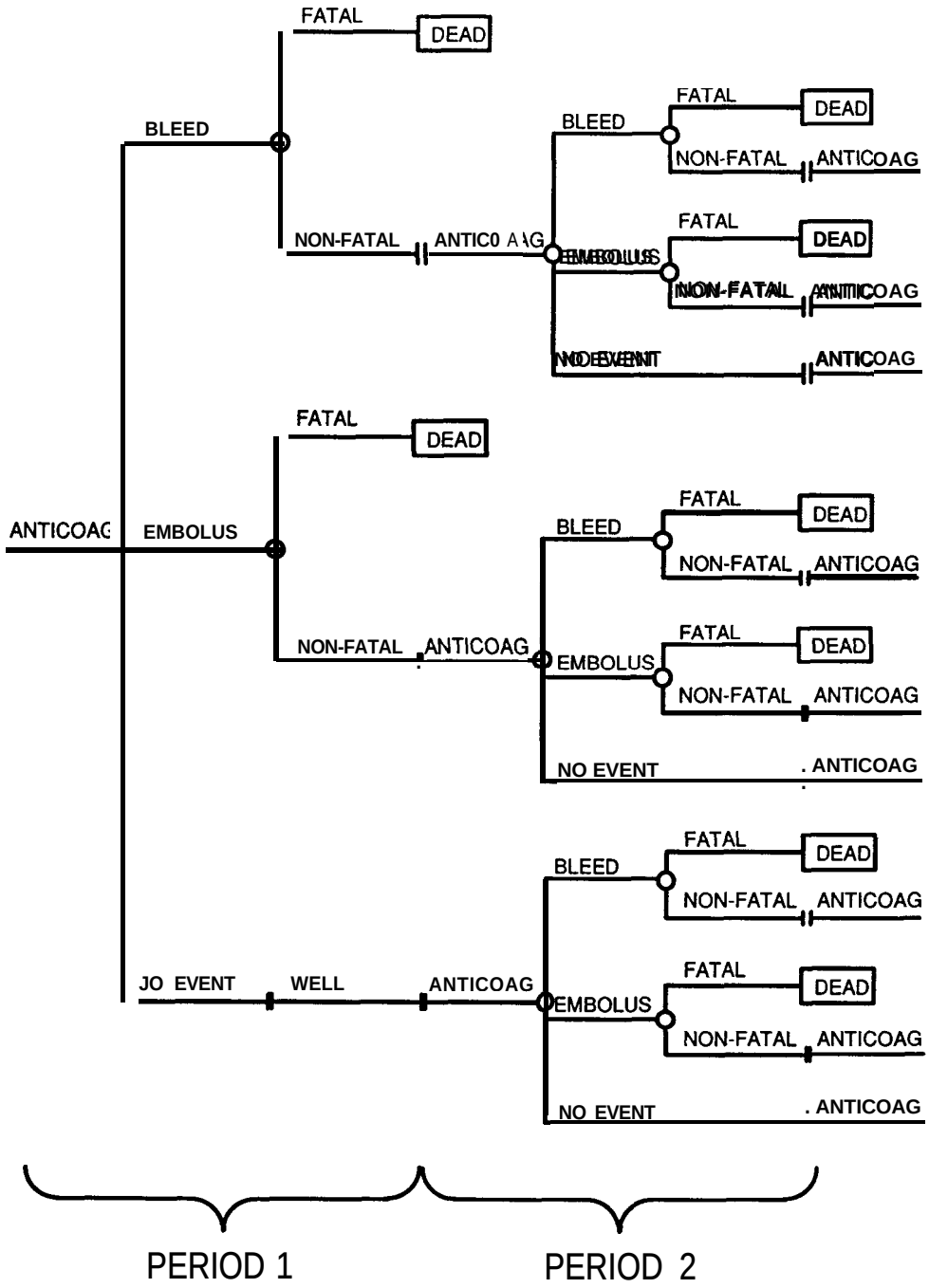
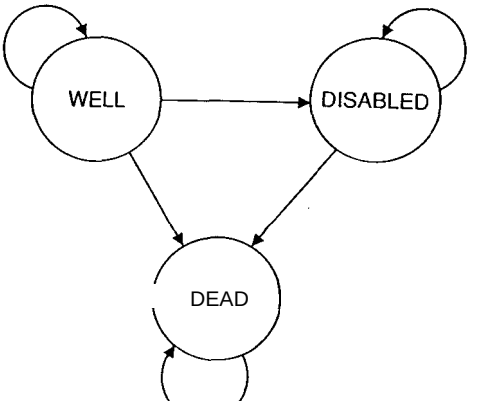


FIGURE 2. **Recursive tree modeling complications of anticoagulant therapy.**

The length of the cycle is chosen to represent a clinically meaningful time interval. For a model that spans the entire life history of a patient and relatively rare events the cycle length can be one year. On the other hand, if the time frame is shorter and models events that may occur much more frequently, the cycle time must be shorter, for example monthly or even weekly. The cycle time also must be shorter if a rate changes rapidly over time. An example is the risk of perioperative myocardial infarction (MI) following previous MI that declines to a stable value over six months.[13] The rapidity of this change in risk dictates a monthly cycle time. Often the choice of a cycle time will be determined by the available probability data. For example, if only yearly probabilities are available, there is little advantage to using a monthly cycle length.

INCREMENTAL UTILITY

Evaluation of a Markov process yields the average number of cycles (or analogously, the average amount of time) spent in each state. Seen another way, the patient is "given credit" for the time spent in each state. If the only attribute of interest is duration of survival, then one need only add together the average



FICURE 3 **Markov-state diagram. Each circle represents a Markov state. Arrows indicate allowed transitions.**

times spent in the individual states to arrive at an expected survival for the process.

$$\text{Expected utility} = \sum_{s=1}^{n} t_s$$

where t_s is the time spent in state s.

Usually, however, the quality of survival is considered important. Each state is associated with a quality factor representing the quality of life in that state relative to perfect health. The utility that is associated with spending one cycle in a particular state is referred to as the ***incremental utility.*** Consider the Markov process depicted in figure 3. If the incremental utility of the DISABLED state is 0.7, then spending the cycle in the DISABLED state contributes 0.7 quality-adjusted cycles to the expected utility. Utility accrued for the entire Markov process is the total number of cycles spent in each state, each multiplied by the incremental utility for that state.

$$\text{Expected utility} = \sum_{s=1}^{n} t_s \times u_s$$

Let us assume that the **DEAD** state has an incremental utility of zero,* and that the WELL state has an incremental utility of 1.0. This means that for every cycle spent in the WELL state the patient is credited with a quantity of utility equal to the duration of a single Markov cycle. If the patient spends, on average, 2.5 cycles in the **WELL** state and 1.25 cycles in the **DISABLED** state before entering the **DEAD** state, the utility assigned would be (2.5 X 1) + (1.25 X 0.7), or 3.9 quality-adjusted cycles. This number is the quality-adjusted life expectancy of the patient.

*** For medical examples, the incremental utility of the absorbing DEAD state must be zero because the patient will spend an infinite amount of time in the DEAD state and if the incremental utility were non-zero, the net utility for the Markov process would be infinite.**

When performing cost-effectiveness analyses, a separate incremental utility may be specified for each state, representing the financial cost of being in that state for one cycle. The model is evaluated separately for cost and survival. Cost-effectiveness ratios are calculated as for a standard decision tree.[10,11]

TYPES OF MARKOV PROCESSES

Markov processes are categorized according to whether the state-transition probabilities are constant over time or not. In the most general type of Markov process, the transition probabilities may change over time. For example, the transition probability for the transition from WELL **to** DEAD consists of two components. The first component is the probability of dying from unrelated causes. In general, this probability changes over time because, as the patient gets older, the probability of dying from unrelated causes will increase continuously. The second component is the probability of suffering a fatal hemorrhage or embolus during the cycle. This may or may not be constant over time.

A special type of Markov process in which the transition probabilities are constant over time is called a Markov chain. if it has an absorbing state, its behavior over time can be determined as an exact solution by simple matrix algebra, as discussed below. The DEALE can be used to derive the constant mortality rates needed to implement a Markov chain. However, the availability of specialized software to evaluate Markov processes and the greater accuracy afforded by age-specific mortality rates have resulted in greater reliance on Markov processes with time-variant probabilities.

The net probability of making a transition from one state to another during a single cycle is called a ***transition probability.*** The Markov process is completely defined by the probability distribution among the starting states and the probabilities for the individual allowed transitions. For a Markov model of n slates, there will be n^2 transition probabilities. When these probabilities are constant with respect to time, they can be represented by an n X n matrix, as shown in table 1. Probabilities representing disallowed transitions will, of course, be zero. This matrix, called the P ***matrix,*** forms the basis for the fundamental matrix solution of Markov chains described in detail by Beck and Pauker.[5]

Table 1 • **P** Matrix

		To		
		WELL	DISABLED	DEAD
From	WELL	**0.6**	**0.2**	**0.2**
	DISABLED	0	0.6	0.4
	DEAD	0	0	**1**

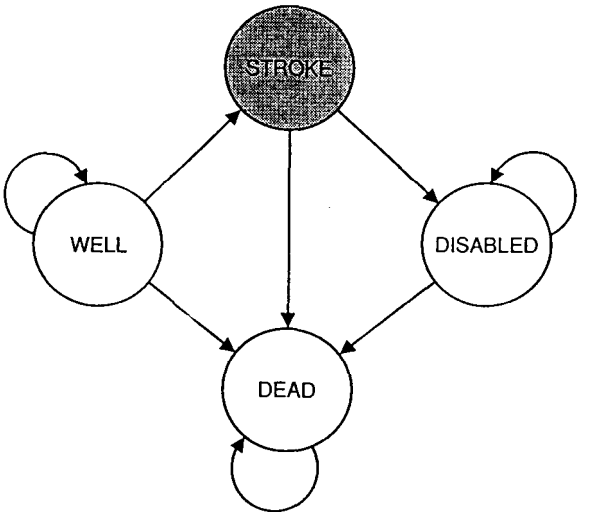


FIGURE 4. Markov-state diagram. The shaded circle labeled "STROKE" represents a temporary state.

THE MARKOV PROPERTY

The model illustrated in figure 3 is compatible with a number of different models collectively referred to as ***finite stochastic processes.*** In order for this model to represent a Markov process, one additional restriction applies. This restriction, sometimes referred to as the ***Markovian assumption***[5] or the ***Markov property,"*** specifies that the behavior of the process subsequent to any cycle depends only on its description in that cycle. That is, the process has no memory for earlier cycles. Thus, in our example, if someone is in the **DISABLED** state after cycle n, we know the probability that he or she will end up in the DEAD state after cycle n + **1.** It does not matter how much time the person spent in the WELL state before becoming DISABLED. Put another way, all patients in the DISABLED state have the same prognosis regardless of their previous histories. For this reason, a separate state must be created for each subset of the cohort that has a distinct utility or prognosis. If we want to assign someone disabled from a bleed a different utility or risk of death than someone disabled from an embolus, we must create two disabled states. The Markovian assumption is not followed strictly in medical problems. However, the assumption is necessary in order to model prognosis with a finite number of states.

MABKOV STATES

In order for a Markov process to terminate, it must have at least one state that the patient cannot leave. Such states are called ***absorbing states*** because, after a sufficient number of cycles have passed, the entire cohort will have been absorbed by those states. In medical examples the absorbing states must represent death because it is the only state a patient cannot leave. There is usually no need for more than one **DEAD** state, because the incremental utility for the DEAD state is zero. However, if one wishes to keep track of the causes of death, then more than one DEAD state may be used.

Temporary states are required whenever there is an event that has only short-term effects. Such states are defined by having transitions only to other states and not to themselves. This guarantees that the patient can spend, at most, one cycle in that state. Figure 4 illustrates a Markov process that is the same as that shown in figure 3 except that a temporary state has been added, labeled STROKE. An arrow leads to STROKE only from the WELL state, and there is no arrow from the STROKE back to itself. This ensures that a patient may spend no more than a single cycle in the STROKE state. Temporary states have two uses. The first use is to apply a utility or cost adjustment specific to the temporary state for a single cycle. The second use is to assign temporarily different transition probabilities. For example, the probability of death may be higher in the STROKE state than in either the WELL state or the **DISABLED State.**

A special arrangement of temporary states consists of an array of temporary states arranged so that each has a transition only to the next. These states are called ***tunnel states*** because they can be visited only in a fixed sequence, analogous to passing through a tunnel. The purpose of an array of tunnel states is to apply to incremental utility or to transition probabilities a temporary adjustment that lasts more than one cycle.

An example of tunnel states is depicted in figure 5. The three tunnel states, shaded and labelled **POST MI1** through **POST MI3,** represent the first three months following an MI. The **POST MI1** state is associated with the highest risk of perioperative death. **POST MI2** and **POST MI3** are associated with successively lower risks of perioperative death. If a patient passes through all three tunnel states without having surgery, he or she enters the POST MI state, in which the risk of perioperative death is constant.

Because of the Markovian assumption, it is not possible for the prognosis of a patient in a given state to depend on events prior to arriving in that state. Often, however, patients in a given state, for example **WELL,** may actually have different prognoses depending on previous events. For example, consider a patient who is WELL but has a history of gallstones. Each cycle, the patient has a certain probability of developing complications from the gallstones. Following a cholecystectomy, the patient will again be **WELL** but no longer has the same probability of developing biliary complications. Thus, the state **WELL** actually contains two distinct populations of people, those with gallstones and those who have had a cholecystectomy. In order for the model to reflect the different prognoses for these two classes of well patients, it must contain two distinct well states, one representing WELL **WITH** GALLSTONES and the other representing WELL, STATUS-POST

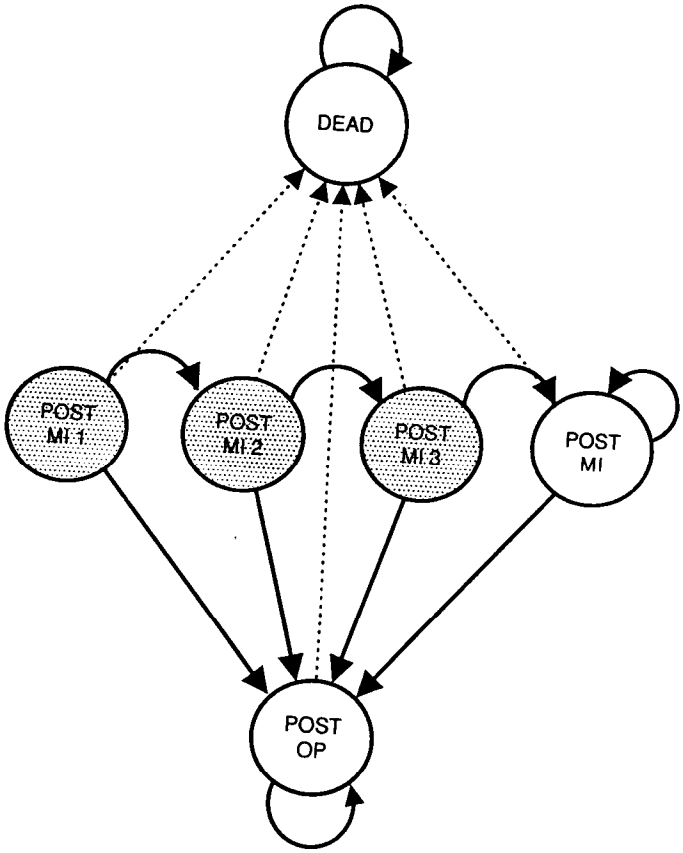


FIGURE 5. Tunnel states: **the three shaded circles represent temporary states that can be visited only in a fixed sequence.**

CHOLECYSTECTOMY. In general, if prognosis depends in any way on past history, it requires that there be one distinct state to represent the different histories.

USE OF THE MARKOV PROCESS IN DECISION ANALYSIS

The Markov process models prognosis for a given patient and thus is analogous to a utility in an ordinary decision tree. For example, if we are trying to choose between surgery and medical therapy, we may construct a decision tree like that shown in figure 6A. In this case, events of interest, such as operative death and cure, are modelled by tree structure "outside" the Markov process. The Markov process is being used simply to calculate survival for a terminal node of the tree. This structure is inefficient, because it requires that an entire Markov process be run for each terminal node, of which there may be dozens or even hundreds. A far more efficient structure is shown in figure 6B. In this case, the Markov process incorporates all events of interest and the decision analysis is reduced simply to comparing the values of two Markov processes. The use of the cycle tree representation (discussed in detail below) permits representing all relevant events within the Markov process.

Representations of Markov Models

THE FUNDAMENTAL MATRIX SOLUTION

When the Markov process has constant transition probabilities (and constant incremental utilities1 for all states, the expected utility may be calculated by matrix algebra to yield the fundamental matrix, which shows, for each starting state, the expected length of time spent in the state. The matrix solution is fast and provides an "exact" solution that is not affected by the cycle length. There are three main disadvantages of the matrix formation. The first is the difficulty in performing matrix inversion. However, this is less of a problem than when Beck and Pauker[5] described the technique, because many commonly available microcomputer spreadsheet programs now perform matrix algebra. The second disadvantage is the restriction to constant transition probabilities. The third disadvantage is the need to represent all the possible ways of making a transition from one state to another as a single transition probability. At least for medical applications, the matrix algebra solution has been largely relegated to the history books. For more details of the matrix algebra solution the reader is referred to Beck and Pauker.[5]

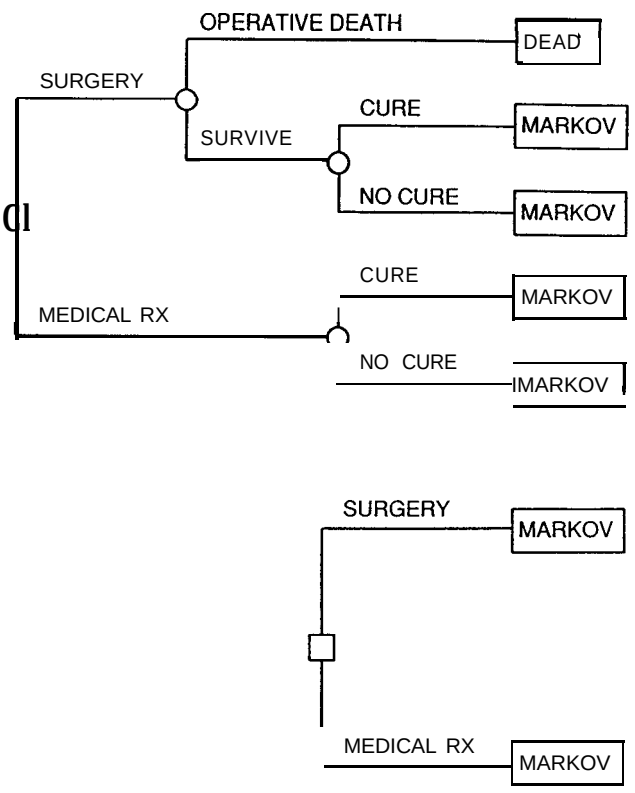


FIGURE 6. **Use of** Markov **processes in a decision model. In panel** *A (top),* **the Markov process is used only as a utility. In panel** *B (bottom),* **the Markov process is used to represent all events.**

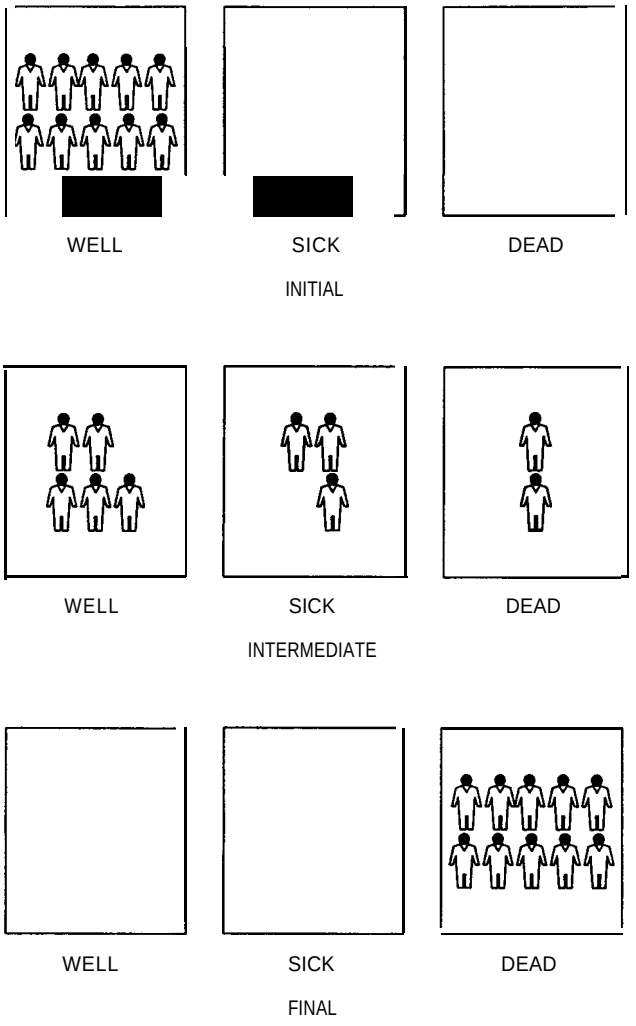


FIGURE 7. **Markov cohort simulation.** Panel A (*top*), **shows the initial distribution with all patients in the** WELL **state. Panel** *B (middle)* **shows the distribution partway through the simulation. Panel C shows the final distribution, with the entire cohort in the** DEAD **state.**

MARKOV COHORT SIMULATION

The Markov cohort simulation is the most intuitive representation of a Markov process. The difference between a cohort simulation and the matrix formulation may be thought of as analogous to the difference between determining the area under a curve by dividing it into blocks and summing their areas versus calculating the area by solving the integral of the function describing the curve. The simulation considers a hypothetical cohort of patients beginning the process with some distribution among the starting states. Consider again the prognosis of a patient who has a prosthetic heart valve, represented by the Markov-state diagram in figure 3. Figure 7A illustrates the cohort at the beginning of the simulation. In this example, all patients are in the WELL state. However, it is not necessary to have all patients in the same state at the beginning of the simulation. For example, if the strategy represents surgery, a fraction of the cohort may begin the simulation in the DEAD state as a result of operative mortality.

The simulation is as follows. For each cycle, the fraction of the cohort initially in each state is partitioned among all states according to the transition probabilities specified by the P matrix. This results in a new distribution of the cohort among the various states for the subsequent cycle. The utility accrued for the cycle is referred to as the ***sum*** and is calculated by the formula:

$$\text{Cycle sum} = \sum_{s=1}^{n} f_s \times U_s$$

where n is the number of states, f_s is the fraction of the cohort in state s, and U_s is the incremental utility of state s. The cycle sum added to a running total that is referred to as the ***cumulative utility.*** Figure shows the distribution of the cohort after a few cycles. Fifty percent of the cohort remains in the **WELL** state. Thirty percent of the cohort is in the **SICK** state and 20% in the **DEAD** state. The simulation is run for enough cycles so that the entire cohort is in the DEAD state (fig. 7C).

The cohort simulation can be represented in tabular form, as shown in table 2. This method may be implemented easily using a microcomputer spreadsheet program. The first row of the table represents the starting distribution. A hypothetical cohort of 10,000 patients begins in the **WELL** state. The second row shows the distribution at the end of the first cycle. In accordance with the transition probabilities specified in the P-matrix (table **1),2,000** patients (20% of the original cohort) have moved to the **DISABLED** state and another 2,000 patients to the DEAD state. This leaves 6,000 (60%) remaining in the **WELL** state. This process is repeated in subsequent cycles. The fifth column in table 2 shows the calculation of the cycle sum, which is the sum of the number of cohort members in each state multiplied by the incremental utility for that state. For example, because the incremental utility of the **DISABLED** state is 0.7, the cycle sum during cycle **1** is equal to 16,000 X **11** + (2,000 X 0.7) = 7,400. The **DEAD** state does not contribute to the cycle sum because its in-

Table 2 • Markov Cohort Simulation

Cycle	WELL	DISABLED	DEAD	Cycle Sum	Cumulative Utility
Start	10,000	0	0	—	—
1	6,000	2,000	2,000	7,400	7,400
2	3,600	2,400	4,000	5,280	12,680
•	•	•	.	.	.
•	•	•	.	.	.
23	0	1	9,999	7	23,752
24	0	0	10,000	<1	23,752
Total	15,000	12,500		23,752	23,752

cremental utility is zero. The sixth column shows the cumulative utility following each cycle.

Because the probabilities of leaving the WELL and DEAD states are finite and the probability of leaving the DEAD sate is zero, more and more of the cohort ends up in the DEAD state. The fraction of the cohort in the DEAD state actually is always less than 100% because, during each cycle, there is a finite probability of a patient's remaining alive. For this reason, the simulation is stopped when the cycle sum falls below some arbitrarily small threshold (e.g., 1 person-cycle) or when the fraction of the cohort remaining alive falls below a certain amount. In this case, the cycle sum falls below 1 after 24 cycles. The expected utility for this Markov cohort simulation is equal to the cumulative utility when the cohort has been completely absorbed divided by the original size of the cohort. In this case, the expected utility is 23,752/10,000, or 2.3752 quality-adjusted cycles. The unadjusted life expectancy may be found by summing the entries in the columns for the WELL and DISABLED states and dividing by the cohort size. Notice that the cohort memberships at the start do not contribute to these sums. Thus, the cohort members will spend, on average, 1.5 cycles in the WELL state and 1.25 cycles in the DISABLED state, for a net unadjusted life expectancy of 2.75 cycles.

THE HALF-CYCLE CORRECTION

The Markov model assumes that during a single cycle, each patient undergoes no more than one state transition. One way to visualize the Markov process is to imagine that a clock makes one "tick" for each cycle length. At each tick, the distribution of states is adjusted to reflect the transitions made during the preceding cycle. The Markov cohort simulation requires explicit bookkeeping (as illustrated in table 2) during each cycle to give credit according to the fraction of the cohort in each state. In the example illustrated in table 2, the bookkeeping was performed at the end of each cycle.

In reality, transitions occur not only at the clock ticks, but continuously throughout each cycle. Therefore, counting the membership only at the beginning or at the end of the cycle will lead to errors. The process of carrying out a Markov simulation is analogous to calculating expected survival that is equal to the area under a survival curve. Figure 8 shows a survival curve for members of a state. The smoothness of the curve reflects the continuous nature of state transitions. Each rectangle under the curve represents the accounting of the cohort membership during one cycle when the count is performed at the end of each cycle. The area of the rectangles consistently underestimates the area under the curve. Counting at the beginning of each cycle, as in figure 9, consistently overestimates survival. To more accurately reflect the continuous nature of the state transitions, we make the assump-

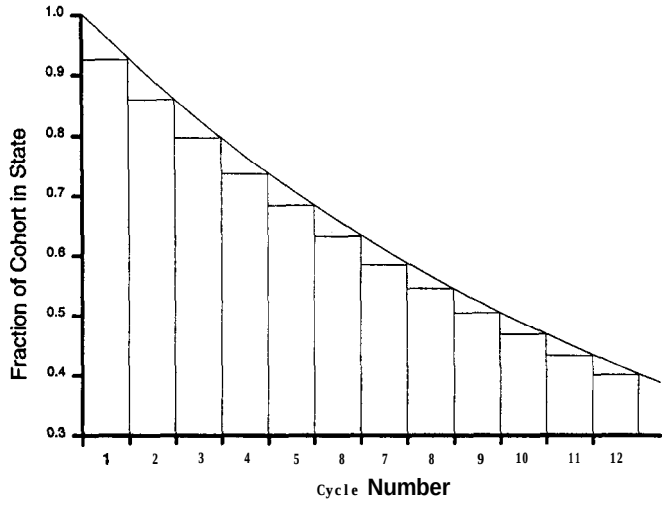


FIGURE 8. **Counting cohort membership at the end of each cycle.**

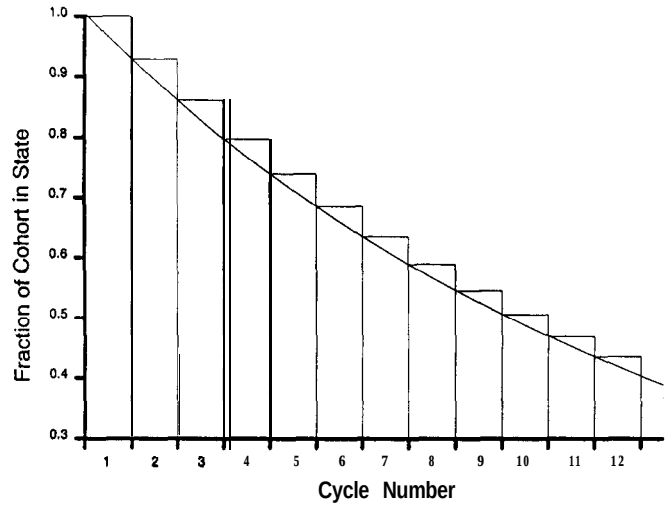


FIGURE 9. **Counting cohort membership at the beginning of each cycle.**

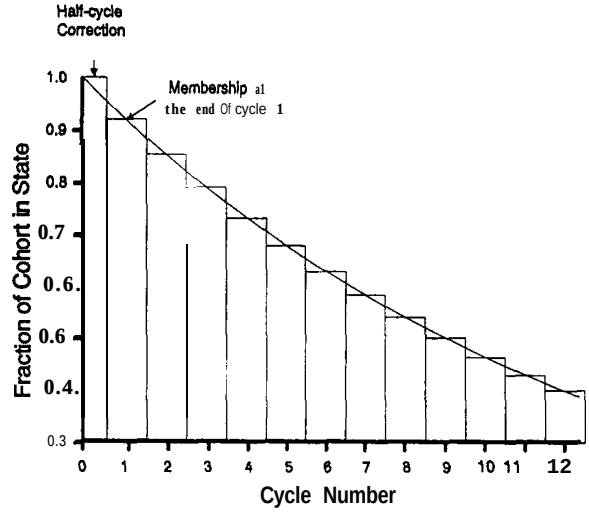


FIGURE 10. **Illustration of the half-cycle correction.**

tion that state transitions occur, on average, ***halfway through*** each cycle. There is no way to determine the state membership in the middle of the cycle. However, if we consider the count at the end of each cycle to be in the middle of a cycle that begins halfway through the previous cycle and ends halfway through the subsequent cycle, as in figure **10,** then the under- and overestimations will be balanced. This is equivalent to shifting all cycles one half cycle to the right. We must then add a half cycle for the starting membership at the beginning to compensate for this shift to the right. Adding a half cycle for the example in table 2 results in an expected utility of 2.875 quality-adjusted cycles and a life expectancy of 3.25 cycles.

The shift to the right makes no difference at the end of the simulation if the cohort is completely absorbed because the state membership at that time is infinitesimal. However, if the simulation is terminated prior to the absorption of the cohort, the shift to the right will result in overestimation of the expected survival. Therefore, for simulations that terminate prior to absorption, an additional correction must be made by subtracting a half cycle for members of the state who are still alive at the end of the simulation. The importance of the half cycle correction depends on cycle length. If the cycle length is very short relative to average survival, the difference between actual survival and simulated survival (as shown in figure 8) will be small. If the cycle time is larger relative to survival, the difference will be more significant. The interested reader should note that the fundamental matrix representation is equivalent to counting state membership at the ***beginning*** of each cycle. Therefore, the correction that should be applied to the result of a matrix solution is ***subtraction*** of one half cycle from the membership of each starting state.

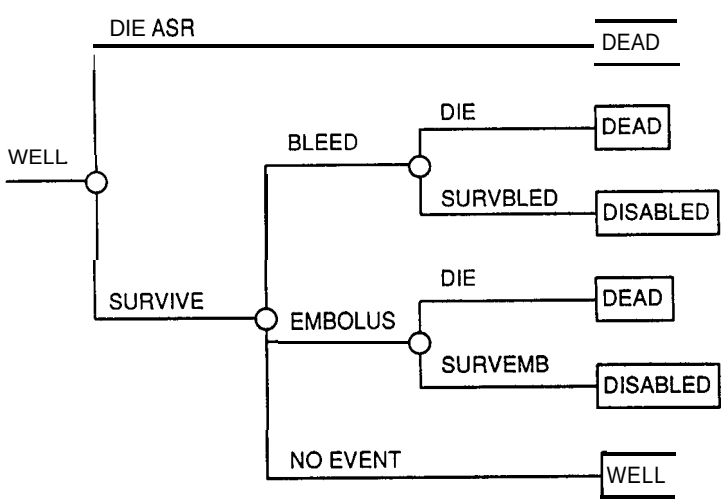


FIGURE 11. **Probability tree corresponding to the WELL state**

THE MARKOV-CYCLE TREE

In the preceding discussion, transition probabilities were provided as if they were elemental data supplied with a problem. However, for actual clinical settings, transition probabilities may be quite complicated to calculate because transitions from one state to another may happen in a variety of ways. For example, a patient in the WELL state may make a transition to the **DEAD** state by having a fatal stroke, by having an accident, or by dying of complications of a coexisting disease. Each transition probability must take into account all of these transition paths. Hollenberg[15] devised an elegant representation of Markov processes in which the possible events taking place during each cycle are represented by a probability tree.

The probability tree corresponding to the WELL state is illustrated in figure 11. It contains a chance node modeling the occurrence of death from age, sex, and race (ASR)-specific mortality, the branch labelled DIE ASR. If the patient does not die from natural causes, the branch labelled SURVIVE leads to a chance node modelling whether the patient has a BLEED or an **EMBOLUS,** either of which may be fatal. If neither BLEED nor **EMBOLUS** occurs (the branch NO EVENT), the patient remains WELL. Each terminal node in the probability tree is labelled with the name of the state in which a patient reaching that terminal node will begin the next cycle. Thus, a patient reaching any terminal node labelled DEAD will begin the next cycle in the DEAD state. A patient surviving either an embolus or a bleed will begin the next cycle in the DISABLED state. The probability tree for patients beginning in the DISABLED state is identical to that for the WELL state, except that patients having NO EVENT will still be DISABLED. The probability tree for patients beginning in the DEAD state consists only of the terminal node labelled with the name of the DEAD state since no event is possible, and a patient in the DEAD state will always remain in that state.

The subtrees are attached to a special type of node designated a Markov node as depicted in figure 12. There is one branch of the Markov **node** for each Markov state. Each probability from the Markov node to one of its branches is equal to the probability that the patient will ***start*** in the corresponding state. The Markov node together with its attached subtrees is referred to as a Markov-cycle tree[15] and, along with the incremental utilities and the probabilities of the branches of chance nodes, is a complete representation of a Markov process. Starting at any state branch, the sum of the probabilities of all paths leading to terminal nodes labelled with the name of a particular ending state is equal to the transition probability from the beginning state to the ending state. For example, the highlighted paths in figure 12 show all transitions from WELL to **DISABLED.**

EVALUATING CYCLE TREES

A cycle tree may be evaluated as a Markov cohort simulation. First, the starting composition of the cohort is determined by partitioning the cohort among the states according to the probabilities leading from the Markov node to the individual branches. Each subtree is then traced from its root to its termini ("folding forward"), partitioning the subcohort for the corresponding state according to the probability tree. The result is a new distribution of the cohort among the states, which reflects how the cohort appears after a single cycle. The fraction of the cohort currently in each state is then credited with the appropriate incremental utility to form the cycle sum, which is added to the cumulative utility. The new distribution of the cohort is then used as the starting distribution for the next cycle. The process is repeated until some predetermined criterion is reached, usually when the quantity of utility accumulating for each state drops below some specified small quantity. This occurs when the fraction of the cohort in the DEAD state approaches one.

ADVANTAGES OF THE CYCLE TREE REPRESENTATION

Cycle trees have many of the same advantages that decision trees have for modelling complex clinical situations. They allow the analyst to break up a large problem into smaller, more manageable ones. This clarifies issues for the analyst and for others trying to understand the results. The use of subtrees promotes appropriate symmetry among the various states, thus enhancing the fidelity of the model. The model provides a great deal of flexibility when changing or refining a Markov model. If a single component probability or a detail of a subtree needs to be changed, this can be done without recalculating the aggregate transition probabilities. Finally, the disaggregation of transition probabilities permits sensitivity analysis to be performed on any component probability. Because of its advantages, the cycle tree representation has been used most often in recently published Markov decision analyses.[6-9]

MONTE CARLO SIMULATION

As an alternative to simulating the prognosis of a hypothetical cohort of patients, the Monte Carlo simulation determines the prognoses of a large number of individual patients. This is illustrated in figure 13. Each patient begins in the starting state (i.e., the WELL state), and at the end of each cycle, a random-number generator is used together with the transition probabilities to determine in which state the patient will begin the next cycle. Just as for the cohort simulation, the patient is given credit for each cycle spent in a non-DEAD state and each state may be adjusted for quality of life. When the patient enters the DEAD state, the simulation is stopped. For the example in figure 13, the patient spends two cycles in the WELL state and three cycles in the DISABLED state before being "absorbed," resulting in a utility of (2 X 1) + (3 X 0.7) or 4.1 quality-adjusted cycles. The process is repeated a very large number (on the order of 10^4) of times. Each trial generates a quality-adjusted survival time. After a large number of trials, these constitute a *distribution* of survival values. The mean value of this distribution will be similar to the expected utility obtained by a cohort simulation. However, in addition to the mean survival, statistical measures such as variance and standard deviation of the expected utility may be determined from this distribution. It should be noted that a Markov cycle tree may be evaluated as a Monte Carlo simulation.

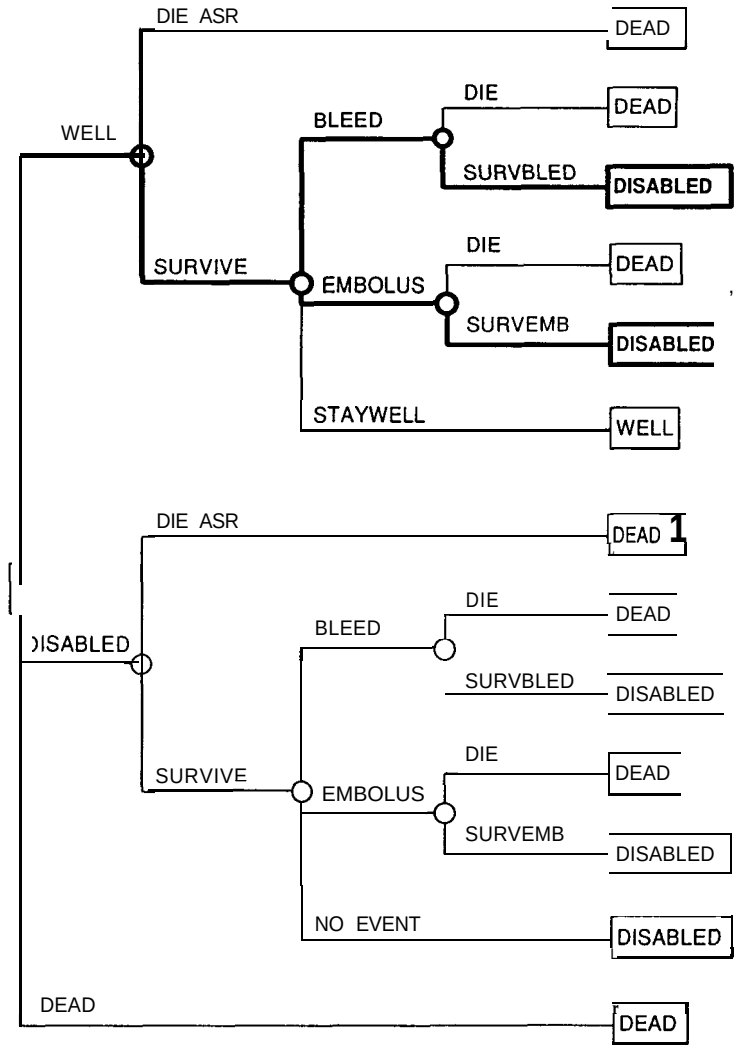


FIGURE 12. Complete Markov-cycle tree corresponding to the anticoagulation problem.

that has experienced the event is equal to $1 - f$. Thus, the curve describing the probability that the event will occur in time t is simply $1 - f$, or $1 - e^{-rt}$ as shown in equation 1. The probability of transition in time t is always less than the corresponding rate per time t because as the cohort members die, fewer are at risk for the transition later in the time period. When the rate is small or t is short, the rate and probability are very similar. Often, data supplied for an analysis provide rates of complications. For use in a Markov analysis, these rates must be converted to the corresponding transition probabilities by substituting the Markov-cycle length for t in equation 1.

PRECAUTIONS IN CHANGING THE CYCLE LENGTH

When changing the Markov-cycle duration from yearly to monthly, one cannot simply divide the calculated transition probabilities by 12 to arrive at the appropriate transition probabilities for the shorter cycle. If the original rate is a yearly rate, then the monthly probability is $p = 1 - e^{-r/12}$. If one has only the yearly transition probability and not the rate, the transition probability can be converted to a rate by solving equation 2 for r:

$$r = -\frac{\ln(1 - p)}{t}$$

Then, the calculated rate is used, as above, to recalculate the transition probability.

TIME DEPENDENCE OF PROBABILITIES

In the most general case, the transition probabilities in a Markov model vary with time. An obvious example is the probability of death, which increases as the cohort ages. If the time horizon for the analysis is a long one, the mortality rate will increase significantly during later cycles. There are two ways of handling such changing probabilities. One is with a continuous function, such as the Gompertz function[3] For each clock cycle, the appropriate mortality rate is calculated from a formula and converted to a transition probability.

Some rates are not easily described as a simple function. One example is the actual mortality rate over a lifetime, which initially is high during early childhood, falls to a minimum during late childhood, and then gradually increases during adulthood. Another example is the risk of acquiring a disease (such as Hodgkins' disease) that has a bimodal age distribution. In such cases, the necessary rates (or corresponding probabilities) may be stored in a table, indexed by cycle number, and retrieved as the Markov model is evaluated. Some computer software used for evaluating Markov processes provides facilities for constructing and using such tables.

DISCOUNTING: TIME DEPENDENCE OF UTILITIES

Incremental utilities, like transition probabilities, may vary with time. One important application of this time dependence is the discounting used in cost-effectiveness analyses.[10] This is based on the fact that costs or benefits occurring immediately are valued more highly than those occurring in the future. The discounting formula is:

$$U_t = \frac{U_0}{(1 + d)^t}$$

where U_t is the increment utility at time t, U_0 is the initial incremental utility, and d is the discount rate.[10] Because of the time variance, discounting cannot be used when the fundamental matrix solution is used.

A Mailed Example

The following example is a Markov implementation of a decision analysis that has been published in detail elsewhere as an ordinary decision tree.[16,17] This analysis was performed for an actual patient at the New England Medical Center. The implementation of the model is a Markov-cycle tree as used by two specific decision analysis microcomputer programs Decision Maker'" and SMLTREE.[19]

Case history. A 42-year-old man had had a cadaveric kidney transplant **18** months previously and had done well except for an early rejection episode, which had been treated successfully. He had maintained normal kidney function. While he was receiving standard treatment with azathioprine and prednisone, however, two synchronous malignant melanomas appeared and required wide resection. Continuation of immunosuppressive therapy increases the chance of another, possibly lethal melanoma. Cessation of this therapy ensures that the patient's kidney will be rejected and will require his return to dialysis, a therapeutic modality he prefers to avoid.

The key assumptions in the construction of this model are:

1. If therapy is stopped, the patient will reject the kidney immediately.
2. If therapy is continued, the patient still may reject the kidney, but with a lower probability.
3. If the patient rejects the kidney despite continuation of therapy, the therapy will be stopped at the time the rejection occurs.
4. A second transplant will not be considered.
5. Quality of life is lower on dialysis than with a functioning transplant. Based on the original utility assessment from the patient, the utility of life on dialysis was 0.7 and that of life with a functioning transplant **1.0.'"**

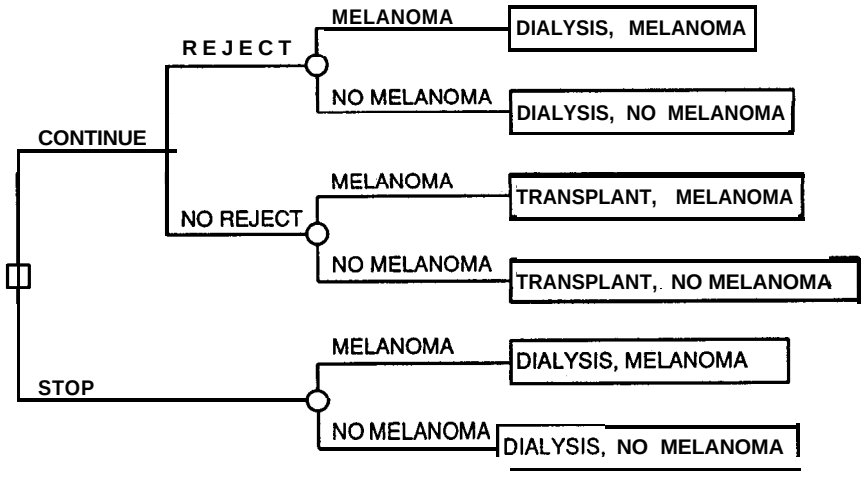


FIGURE 14. **Simple decision tree for the kidney transplant, melanoma case.**

6. **No adjustment is made to quality of life for having** recurrent melanoma.

7. The patient's life expectancy is reduced because of having had a renal transplant and a previous melanoma. It will be further reduced if the patient goes on dialysis or if melanoma recurs.

The simple tree modeling this problem is shown in figure **14.** There are two branches of the decision node, representing **CONTINUE** and **STOP** therapy, respectively. In the case of CONTINUE, a chance node models the occurrence of **REJECT or NO REJECT.** The development of a new melanoma is modelled by the chance node with the branches **MELANOMA** and **NO MELANOMA.** Terminal nodes represent the six possible combinations of therapy, renal status, and occurrence of a new melanoma. The two combinations representing STOP **THERAPY** and NO **REJECT** are assumed not to occur. Probabilities in this model must be assigned to reflect the different risks of developing a new melanoma depending on whether or not therapy has been continued. The lowest probability is for patients whose therapy is stopped immediately. The highest probability is for those whose therapy is continued indefinitely. Because therapy will be stopped, patients who experience rejection after an initial period of continuing therapy will have an intermediate risk of melanoma recurrence.

Because it was a simple tree, the original model required several simplifying assumptions. The first was that recurrent melanoma occurred at a fixed time in the future (one year)'" although, in reality, it may occur at any time. The second was that transplant rejection occurred at a fixed time, the midpoint of the patient's life expectancy. Therefore, the utility of continuing therapy, then experiencing transplant rejection was assigned the average of the utilities for transplant and dialysis. If the patient values time on dialysis differently now compared with later, this is an oversimplification. The third assumption was that the probability of melanoma recurrence in this intermediate scenario was the average of the high and low probabilities. Again, this is an oversimplification because the patient actually has a high probability while on the therapy and a low probability while off it. The Markov model can address all of these issues.

The Markov decision model is shown in figure **15** and figure **16.** The root of the tree in figure 15 is a decision node with two branches representing the two choices **CONTINUE** and STOP. The Markov-cycle tree depicted in figure **16** consists of a Markov node with one branch for each Markov state. If we assume that the utility of a state depends only on whether the patient is on dialysis or not and that the probability of melanoma depends only on whether or not the patient is receiving immunosuppressive therapy, then only five states are required to represent the scenario. These are (from top to bottom) **TRANSWEL** (transplant well), **TRANSMEL** (transplant with melanoma), **DIALWELL** (dialysis, no melanoma), **DIALMEL** (dialysis and melanoma), and **DEAD.** Separate states are not needed based on treatment because it is assumed that patients in the transplant states are on immunosuppressive therapy and those in the dialysis states are not.

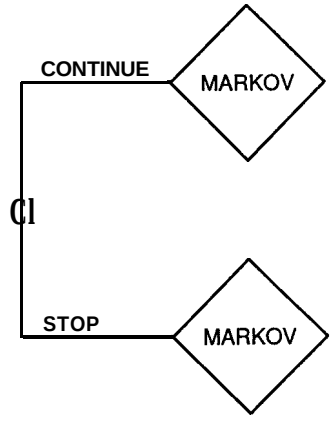


FIGURE 15. **Root of the tree representing the Markov model of the kidney transplant, melanoma case.**

INITIAL PROBABILITIES

The first task is to assign probabilities to the branches of the Markov node. Recall that these probabilities represent the probabilities of starting in the individual states. For the **CONTINUE** strategy, all patients begin in the **TRANSWEL** state, so the probability of that state should be **1.** Similarly, for the **STOP** strategy, all patients begin in the **DIALWELL** state. We can implement these assumptions by assigning the probabilities of the **TRANSWELL** and **DIALWELL** branches as variables. A binding set between the strategy branch and the Markov node can set the appropriate variable to **1.** Thus, the same Markov-cycle tree can be used as a subtree'" to represent the prognosis for both strategies.

SUBSEQUENTPROGNOSIS

Each branch of the Markov node is attached to a subtree that models the possible events for each Markov state. The most complex is for the **TRANSWEL** state, shown at the top of figure 16. The first event modelled is the chance of dying from all causes (the branch Die). Die leads to a terminal node. In this case the utility is **DEAD,** because a patient who dies during one cycle will begin the next cycle in the **DEAD** state. For patients who do not die (the branch Survive), the next chance node models the chance of transplant rejection (the branches Reject and NoReject). Subsequent to each of these branches is a chance node modelling the risk of recurrent melanoma (the branches Recur and NoRecur and Recur2 and NoRecur2) Each of these branches leads to a terminal node. For Recur following Reject, the appropriate state is **DIALMEL,** for Recur2 following NoReject the appropriate sate is **TRANSMEL .** For NoRecur following Reject and NoRecur2 following NoReject, the appropriate states are **DIALWELL** and **TRANSWEL,** respectively. Only the latter branch represents a return to the starting state.

The event tree for the **TRANSMEL** state is also shown in figure 16. It is simpler than that for **TRANSWEL** because the risk of melanoma recurrence need not be modeled. Assignment of terminal states is similar to that for the **TRANSWEL** state except that patients may not make a transition to the **TRANSWEL** or **DIALWELL** state. Similarly, the probability tree for the **DIALWELL** state models only the risks of death and of melanoma recurrence and that for the **DIALMEL** state models only the risk of death. The event tree for the **DEAD** state is simply a terminal node assigned to the state **DEAD,** since no event is possible and all patients return to the **DEAD** state in the subsequent cycle.

CHOICE OF CYCLE LENGTH

Before the probabilities can be assigned, the analyst must decide on the cycle length. The cycle length should be short enough so that events that change over time can be represented by changes in successive

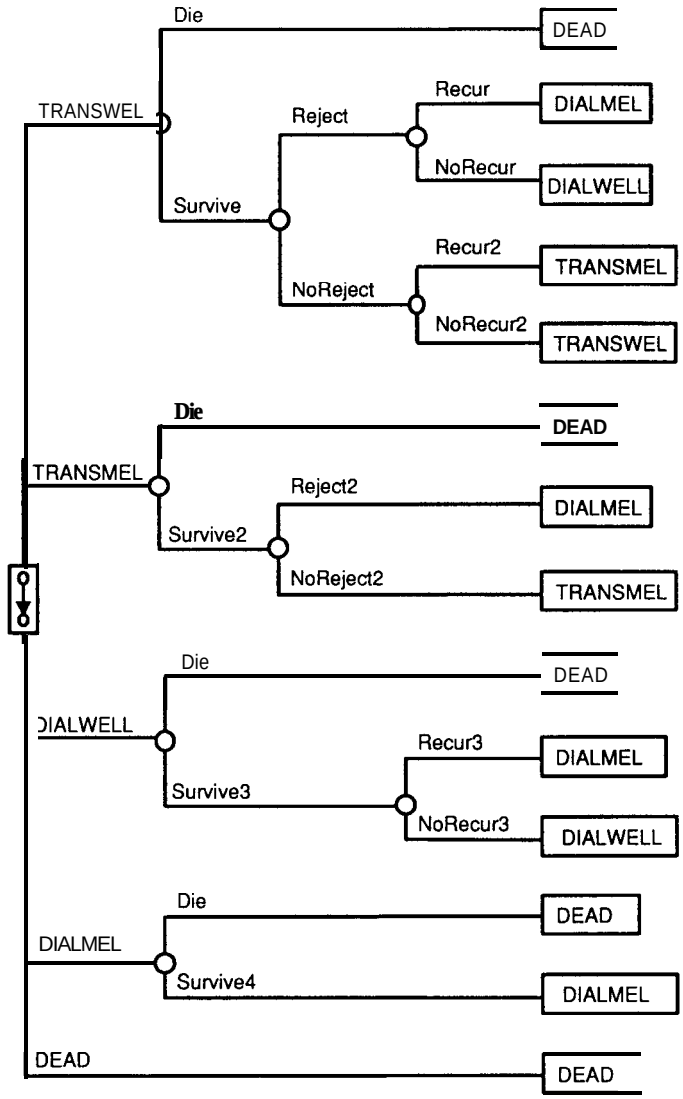


FIGURE 16. Markov-cycle tree representing the kidney transplant, melanoma case.

cycles. For example, if the risk of allograft rejection were markedly different in month 3 than in month **1,** then a monthly cycle should be used. Another consideration is that the cohort simulation is an approximation and will more closely approximate the "exact" solution when the cycle length is short. In practice, however, it makes little difference whether the cycle length is one year or one month, if the appropriate half-cycle corrections are made.[20] A final consideration is evaluation time. A monthly cycle length will result in a 12-fold increase in evaluation time over a yearly cycle length. For this example, since there is no important change during a year, a yearly cycle length is used.

ASSIGNMENT OF PROBABILITIES

The next task is to assign probabilities to the events in each event tree. Each state has a chance node representing the occurrence of death during a given cycle.

This probability is based on the mortality rate for each state, which consists of two components, the baseline mortality rate and the excess mortality due to any comorbid diseases. The baseline mortality rate (mASR) depends on the patient's age, sex, and race. We can assign a different mASR for each cycle to reflect the increasing mortality rate as patients get older. The most convenient way to implement this is to use a table of age-specific mortality rates and look up the appropriate value for each cycle. With a yearly cycle length, the patient's age at the end of cycle n is:

$$Age_{cycle\ n} = StartAge + n$$

There are three refinements to the age used to calculate the cycle-specific age. First, because we assume that transitions occur, on average, halfway through a cycle, the age should be reduced by 0.5 cycle. Second, published mortality rates for patients of a given age (e.g., 50years old) are derived from all patients between that age and the next (50 and **51** years). Thus, the published rate for age 50 actually applies to a group with an average age of 50.5 years and the cycle-specific age should be reduced by an additional 0.5 year to retrieve the appropriate rate. Finally, deaths are slightly more likely to occur among the older members of a heterogeneous cohort and toward the end of a year (when all members are older), so that the observed death rate applies to patients who are slightly older than the average. Empirically, reducing the age by an additional 0.1 to 0.2 years corrects for these effects. Thus, the starting age should be corrected according to the formula:

$$StartAge := StartAge - 0.65 - 0.5 \times cyclen$$

where cyclen is the length of the Markov cycle in **years.** For a detailed discussion of these corrections, the interested reader is referred to Sonnenberg and Wong.[20]

The mortality rate may be retrieved from the table (MTABLE) by the following expression, where StartAge is corrected as above and m.CYCLE is the Markov-cycle number:

$$mASR := MTABLE[StartAge + m.CYCLE \cdot cyclen]$$

For this example, the initial value of mASR is 0.00361 year (for a 43-year-old male). The excess component of mortality due to the patient's coexisting diseases is added to the baseline mortality rate to produce a total compound mortality rate.' The total mortality rate may then be used to calculate the probability of death during any cycle:

$$pDIE := 1 - Exp(-MTABLE[StartAge + m.CYCLE \cdot cyclenl + mEXCESS)$$

DEFERRED EVALUATION

There are two ways to introduce the expression defining the probability of death into the Markov decision model. At first glance, it may seem that the expression above can be placed in a binding proximal to the Markov node, since it is shared by all branches. However, the values of mEXCESS are different for the individual states. Moreover, the value of the expression for pDIE should change for each cycle of the simulation as mASR increases. A simple binding would be evaluated only once and thus would not allow the value of the expression to change. One solution is to place a binding with the above expression on each branch of the Markov node. However, this is cumbersome, because it requires entering the expression four times (it isn't needed for the DEAD state), and slows evaluation, because the entire expression must be placed on the binding stack during the evaluation of each state during each cycle.

An ideal solution is provided by deferring the evaluation of the expression until the value of pDIE is needed, thus ensuring that the evaluation will use the current value of m.CYCLE and the appropriate local value of mEXCESS. This is accomplished using the PASSBIND function in Decision Maker[18] and SMLTREE.[19] This function tells the computer program to place the *entire expression* on the binding stack instead of evaluating the expression first. Thus, when the value of pDIE is needed at any time, anywhere in the tree, the expression will be evaluated with the prevailing values of m.CYCLE and mEXCESS. The expression thus can be entered in a binding proximal to the Markov node. The required binding expression is:

$$pDIE := PASSBIND(1 - EXP(-(MTABL[StartAge + \mathbf{m.CYCLE} * cyclen] + mEXCESS)))$$

Each branch of the Markov node then needs a binding for the appropriate value of mEXCESS. The values of the probabilities of pReject and pRecur depend on whether the patient is on immunosuppressive therapy and therefore must be specified for each state. The values of mEXCESS, pReject, and pRecur for each state are shown in table 4.

ASSIGNING UTILITIES

As described above, utilities in a Markov cohort simulation are associated with a state, rather than with terminal nodes of the tree. Therefore, each state must be assigned an incremental utility that reflects the value of being in that state for one cycle. In Decision Maker[18] and SMLTREE [19] this is accomplished by setting the values of three' special variables. The variable m.uINCR represents the incremental utility of a state for one cycle. The variable m.uINIT is a one-time adjustment to the incremental utility that is made at the beginning of the Markov simulation. It is used to im-

Table4 • Mortality Rates and Probabilities

State	mEXCESS	pReject	pRecur
TRANSPLANT WELL	0.054	0.034	0.44
TRANSPLANT MELANOMA	0.153	0.034	n/a
DIALYSIS WELL	0.110	n/a	0.06
DIALYSIS MELANOMA	0.209	n/a	n/a

plement the half-cycle correction and therefore its value is usually set to **0.5 X** m.uINCR. The variable m.uTAIL is used when the Markov cohort simulation is terminated before the entire cohort is in the absorbing state. Its value is added to the incremental utility for a state at the end of the simulation. The tail utility has two uses. One is to represent the prognosis beyond the stopping point in the Markov simulation. For example, if a Markov process is used only to represent the events taking place during the first six months following an operation, then m.uTAIL will represent the life expectancy of the patient beyond the first six months. The second use is to apply the half-cycle correction to a simulation that is stopped prior to absorption of the cohort, even if the subsequent prognosis is of no interest. In this case, the tail utility must be set to -0.5 X m.uINCR.

The values of these special variables are set with bindings on each branch of the Markov node. For the **TRANSWEL** and **TRANSMEL** states the bindings are:

m.uINIT := 0.5
m.uINCR := **1**
m.uTAIL : = 0

The value of m.uINCR is **1** because there is no utility decrement for the **TRANSPLANT** states. m.u.TAIL is 0 because we are planning to run the simulation until the cohort is completely absorbed.

For the **DIALWELL** and **DIALMEL** states, the bindings are:

m.uINIT := 0.35
m.uINCR := 0.7
m.uTAIL := 0

because the **DIALYSIS** states are associated with a utility of only 0.7 relative to perfect health.[16]

For the DEAD state, the bindings are:

m.uINIT := 0
m.uINCR := 0
muTAIL := 0

because no utility accrues for membership in the DEAD state. In practice, these bindings may be omitted for the DEAD state because if their values are not specified, they will be assumed to be zero.

MARKOV-STATE BINDINGS: A REFINEMENT

Examination of figure 16 reveals that a chance node with branches Reject and NoReject appears in two places in the tree. Similarly, a chance node with branches Recur and NoRecur appears in three places in the tree. We would like to use a common subtree to represent these events in all portions of the tree. The problem is that several of the branches are terminal nodes and their utilities apply only to one specific context. The solution to this problem is the use of Markov-state bindings. When Markov-state names are used on the right side of a binding expression, the program substitutes the state on the right side for the variable on the left side wherever it appears. This permits representing the prognoses of all four non-dead states with a single subtree, as in figure 17. With the state bindings shown, this Markov-cycle tree will be functionally identical to the one in figure **16.**

EVALUATION

When the Markov model is evaluated as a cohort simulation, the expected utilities are:

CONTINUE **THERAPY**	**7.4**
STOPTHERAPY	**5.2**

Thus, the analysis favors continuing therapy by a large margin, more than two quality-adjusted life years. If the quality adjustment is removed from the analysis (by setting quality of life on dialysis to unity), then the results are:

CONTINUE THERAPY	**7.8**
STOPTHERAPY	7.5

Conclusion

Markov models consider a patient to be in one of a finite number of discrete states of health. All clinically important events are modelled as transitions from one state to another. Markov processes may be represented by a cohort simulation (one trial, multiple subjects), by a Monte Carlo simulation (many trials, a single subject for each), or by a matrix algebra solution. The matrix algebra solution requires the least computation, but can be used only when transition probabilities are constant, a special case of the Markov process called a Markov chain. The Markov-cycle tree is a formalism that combines the modelling power of the Markov process with the clarity and convenience of a decision-tree representation. Specialized computer software[18,19] has been developed to implement Markov-cycle trees.

The assignment of quality adjustments to incremental utility permits Markov analyses to yield quality-adjusted life expectancy. Discounting may be applied to incremental utilities in cost-effectiveness analyses. The Markov model provides a means of modelling clinical problems in which risk is continuous over time, in which events may occur more than once, and when the utility of an outcome depends on when it occurs.

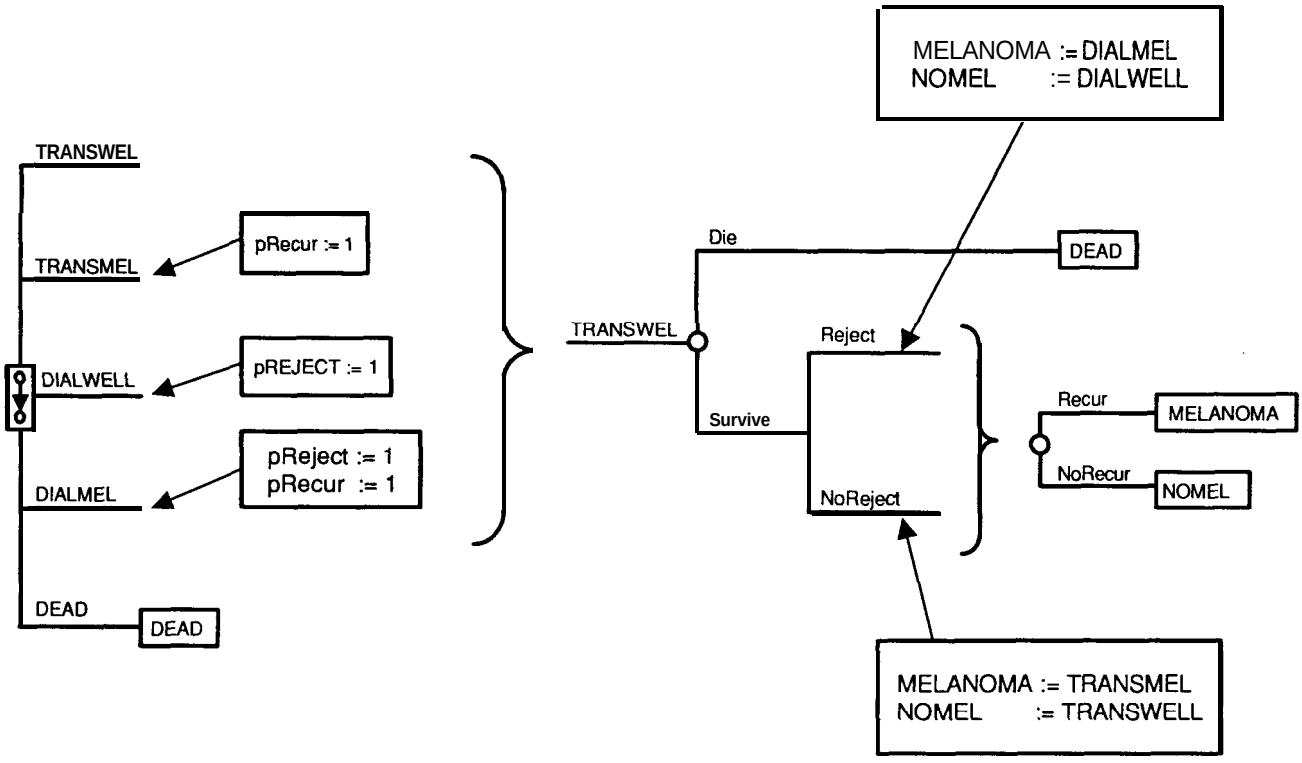


FIGURE 17. Markov-cycle tree using subtrees and state bindings.

Most analytic problems involve at least one of these considerations. Modelling such problems with conventional decision trees may require unrealistic or unjustified simplifying assumptions and may be computationally intractable. Thus, the use of Markov models has the potential to permit the development of decision models that more faithfully represent clinical problems.

References

1. Mehrez A, Gafni A. Quality-adjusted life years, utility theory, and healthy-years equivalents. Med Decis Making 1989;9:142-9.
2. Beck JR, Kassirer JP, Pauker SG. A convenient approximation of life expectancy (the "DEALE"). I. Validation of the model. Am J Med. 1982;73:883-8.
3. Gompertz B. On the nature of the function expressive of the law of human mortality. Philos Trans R Soc London. 1825;115:513-85.
4. National Center for Health Statistics. Vital Statistics of the United States, 1988. Vol II, Mortality, Part A, Section 6. Washington, Public Health Service, 1991.
5. Beck JR, Pauker SG. The Markov process in medical prognosis. Med Decis Making. 1983;3:419-58.
6. Eckman MH, Beshansky JR, Durand-Zaleskl I, Levine HJ, Pauker SG. Anticoagulation for noncardiac procedures in patients with prosthetic heart valves. Does low risk mean high cost? JAMA, 1990;263:1513-21.
7. Wong JB, Sonnenberg FA, Salem D, Pauker SG. Myocardial revascularization for chronic stable angina: an analysis of the role of percutaneous transluminal coronary angioplasty based on data available in 1989. Ann Intern Med. 1990;113:852-71.
8. Hillner BE, Smith TJ, Desch CE. Efficacy and cost-effectiveness of autologous bone marrow transplantation in metastatic breast cancer. Estimates using decision analysis while awaiting clinical trial results. JAMA. 1992;267:2055-61.
9. Birkmeyer JD, Martin CA, O'Connor G-T. Should patients with Bjork-Shiley valves undergo prophylactic replacement? Lancet. 1992;340:520-3.
10. Weinstein MC, Stason WB. Foundations of cost-effectiveness analysis for health and medical practices. N Engl J Med. 1977;296:716-21.
11. Detsky AS, Naglie IG. A clinician's guide to cost-effectiveness analysis. Ann Intern Med. 1990;113:147-54.
12. Lau J, Kassirer JP, Pauker SG. DECISION MAKER 3.0: improved decision analysis by personal computer. Med Decis Making. 1983;3:39-43.
13. Goldman L. Cardiac risks and complications of noncardiac surgery. Ann Intern Med. 1983;98:.504-13.
14. Kemeny JB, Snell JL. Finite Markov Chains. New York: Springer-Verlag, 1976.
15. Hollenberg JP. Markov cycle trees: a new representation for complex Markov processes (abstr). Med Decis Making. 1984;4:529.
16. Cucharal GJ, Levey AS, Pauker SG. Kidney Failure or Cancer: Should Immunosuppression Be Continued in a Transplant Patient with Malignant Melanoma? Med Decis Making. 1984;4:82-107.
17. Kassirer JP, Sonnenberg FA. Decision analysis. In Kelley WN, ed. Textbook of Internal Medicine. Philadelphia: J. B. Lippincott, 1988, 1991.
18. Sonnenberg FA, Pauker SG. Decision Maker: an advanced personal computer tool for clinical decision analysis. Proceedings of the Eleventh Annual Symposium on Computer Applications in Medical Care, Washington, DC.: IEEE Computer Society, 1987.
19. Hollenberg J. SMLTREE: The All Purpose Decision Tree Builder. Boston: Pratt Medical Group; 1985.
20. Sonnenberg FA, Wong JB. Fine-tuning Markov models for life-expectancy calculations. Med Decis Making. 1993;13:170-2.

Erratum:

The equation on page 336 that reads:

```
pDIE := PASSBIND(1-EXP(-MTABLE[StartAge + m.CYCLE*cyclen] + mEXCESS))
```

should have an extra set of parenthesis and should read:

```
pDIE := PASSBIND(1-EXP(-(MTABLE[StartAge + m.CYCLE*cyclen] + mEXCESS)))
```