

Markov Process-Based Monte Carlo Simulation: A Tool for Modeling Complex Disease and its Application to the Timing of Liver Transplantation

Mark S. Roberts
Division of General Internal Medicine
New England Deaconess Hospital/Harvard Medical School
1 Autumn Street, Suite 3A
Boston, Massachusetts 02215, U.S.A.

ABSTRACT

This paper outlines the use of simulation as an extension of the basic methodology of medical decision analysis. The motivation for medical decision analysis is described, and the differences between the standard decision analysis methodologies and those of discrete event simulation are discussed. The technique of Monte Carlo evaluation of Markov process-based models is introduced, its similarity to logical network simulation is noted, and the advantages and characteristics of these models are discussed. Preliminary work on one such model being developed to evaluate the optimal timing of a surgical therapy (liver transplantation) in the declining course of a chronic disease (end-stage liver disease) is presented.

1 INTRODUCTION

Understanding the optimal diagnostic or therapeutic strategy in particular clinical situations is becoming significantly more complicated. As investigators find new relationships between risk factors and disease, the possible combinations of treatment choices and sequences become quite large. Randomized controlled trials, the "gold standard" methodology for determining the "best" treatment, are often not appropriate for evaluating decisions among multiple treatment options given diverse clinical presentations, as they would require too many patients to fill multiple treatment arms with multiple strata. Therefore, decision analysis and other non-experimental methodologies have been used to combine data that currently exists concerning a specific clinical problem, and through multiple "what if" questions, make predictions about likely outcomes and optimal strategies. Initially used to analyze relatively specific problems, medical decision analysis has been increasingly used to address larger and more complicated problems, and several investigators have strained the practical limits of standard decision analysis techniques in their attempts to model clinically complicated situations. This paper will

discuss the basic methodologies of medical decision analysis, describe how some of the limits arise, and illustrate how casting these problems as simulation models facilitates their construction and solution.

2 MEDICAL DECISION ANALYSIS

2.1 Introduction to Medical Decision Analysis

Simulation modeling in medical decision analysis developed from somewhat different origins than the simulation modeling of engineering or manufacturing. The basic structure of medical decision analysis arises from the need to make decisions between diagnostic or therapeutic decisions when knowledge concerning outcomes is uncertain, rather than understand the performance of deterministic systems such as production lines, queues, etc., under the influence of some stochastic process such as arrival times, failure rates, etc.

The most common methodology used to solve decision analysis problems is the standard decision tree; a simple version of which is described in Figure 1. The description of choices emanate from a DECISION node; all relevant outcomes follow as branches of CHANCE nodes; the various possible outcomes are described in TERMINAL NODES which must be valued in the same units (usually life expectancy, utility, etc.) The standard solution to these problems is calculated by "averaging out and folding back", which produces the expected value of a particular decision; that branch with the highest expected value of the outcome variable is the clinically "optimal" choice.

2.2 Markov Processes to Describe Disease States

However, standard decision trees have serious limitations in their ability to model complex situations, especially when outcomes or events occur (or may re-occur) over time. To help solve this problem, Beck and Pauker (1983) introduced the use of Markov processes

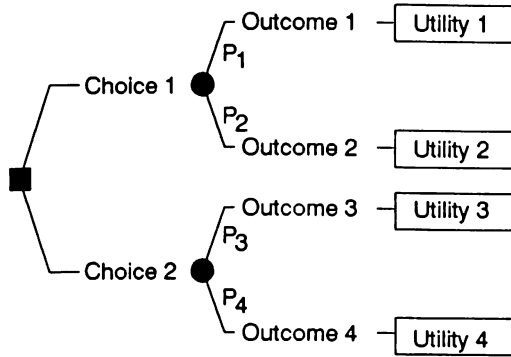


Figure 1: Basic Decision Tree

to model recurrent health states and future events. Because of the ease with which time is incorporated into such models, Markov processes have become quite common in the modeling of several chronic diseases. Standard Markov processes (illustrated in Figure 2) assume that the problem can be fully described as a series of mutually exclusive and all encompassing "states" (WELL, SICK, DEAD); movement between states is controlled by a "transition matrix" (see Table 1), which denotes what portion of the cohort will move from state i to state j in any time period. Evaluation is typically carried out by cohort analysis, where the portion of patients in each state is multiplied by the value of being in that state. These values are summed over all time periods and states, and produce an expected outcome (utility, survival, etc) for the cohort who started the model.

2.3 Limits of Standard Markov Processes, the "Lack of Memory"

Although Markov processes facilitate the construction of models in which components vary over time, they are handicapped by the strict Markovian

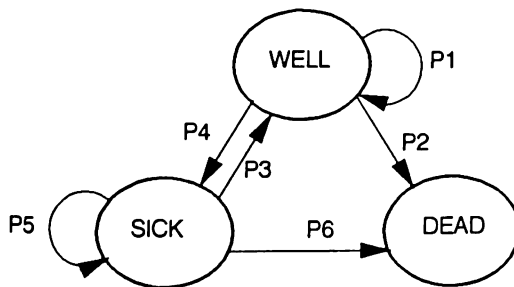


Figure 2: Standard Markov Process

assumption that transition probabilities are "path independent". In their purest form (Markov chains), probabilities may vary only by state, and cannot use information about how or when a particular member of the cohort arrived in that state. Straightforward extensions (Markov cycle trees) use cohort analysis to solve the models, rather than analytic solutions based on the state-to-state transition matrix. This allows for the simple incorporation of time-varying but state-specific variables, such that transition probabilities can be modeled as a function of time (proxied by the cycle number) and the state (Hollenberg, 1984).

Table 1: Transition Matrix for Markov Model in Fig. 2

		Time t+1		
Time t	WELL	SICK	DEAD	
WELL	P1	P4	P2	
SICK	P3	P5	P6	
DEAD	-	-	1	

3 MARKOV PROCESS BASED SIMULATION MODELS

It follows from the above discussion that the major impediment to increasing the complexity of state-based Markov models is the requirement that all of the members of a state be treated the same, regardless of how or when they arrived in a particular state. Since cohort analysis is the most common method for solving Markov processes, each state is in fact composed of a heterogeneous portion of the cohort at any given time. Members of any given state will have arrived there at different times, and may have arrived having followed different paths through the other states. Since standard cohort analysis of a Markov process requires that all members of a state be treated the same (have the same transition probabilities) the heterogenous nature of the portion of the cohort is lost, and all members of the cohort are treated the same, regardless of history. However, if the Markov process is analyzed by Monte Carlo methods (one member of the cohort at a time), the constraint of path independence is removed, and the heterogeneity of members of the cohort arriving at different times is preserved. This form of analysis converts the standard Markov Process into a model of similar genre to traditional simulation modelling.

3.1 Monte-Carlo Analysis of Markov Processes

The evaluation of a Markov Process by Monte Carlo analysis eliminates the "cohort" approach to the solution of a Markov process, and releases the model from the strict assumption of path independence. By simulating each individual member of a the cohort separately, and sending them though the model one at a time in succession; there is only one particular member of the cohort in exactly one state at every point in time. By keeping track of the relevant model occurrences in a global vector, transition probabilities can become functions of those variables, and therefore, functions of the past history and/or path through the process.

When only categorical variables are used, the Monte Carlo evaluated Markov Process may be considered to be "*simulating*" a larger, fully structurally specified Markov process with essentially 2^N states (for dichotomous variables). Figure 3 illustrates a simple example. The left side of the figure describes a simple standard Markov process in which the probability of dying from a disease depends upon whether or not the patient had been previously vaccinated or not. This model requires 5 states, including the all-absorbing DEAD state. The right hand side illustrates the simplicity of structure gained by Monte carlo analysis. Since there is only one member of the cohort in the model at a time, a simple flag can keep track of whether vaccination has occurred or not. The transition probability is then made a function of whether or not the vaccination flag is 0 or 1, rather than be a constant attached to the SICK state. It can be easily shown empirically that the two models give exactly the same results, provided sufficient iterations of the Monte Carlo model are performed.

In essence, there is little if any technical difference

between this type of model and the simulation of logical networks introduced by Roberts in the early 1980's (Roberts and Klein, 1984) and summarized by Dittus and Roberts (1989). The conceptual difference is that the technique described in this paper starts from a methodology (Markov processes) that is now familiar to many medical decision analysts, and with straightforward simulation extensions, creates a state-based, time-varying model that contains all of the advantages of simulation, yet retains the basic Markov cycle tree structure.

3.2 Comparison To Discrete Event Simulation

The major difference between Monte Carlo evaluation of Markov processes and standard discrete event simulation models is that there is no competition for resources, membership in states, or transitions between states as there may be in discrete event simulation modeling. This eliminates the need for structures that represent queues, and limits the need to keep tract of multiple, simultaneous and/or synchronized events. A major strength of discrete simulation modeling for most industrial purposes (that it can incorporated resource constraints, bottlenecks, etc) is rarely used in standard medical decision analyses. Simply stated, the general purpose of medical decision analysis is to examine what the optimal diagnosis or treatment of a specific disease process is, not whether any two given patients could undergo that treatment at the same time. The advantage of using simulation modeling in the current example is that it allows for the construction of a problem that would otherwise be significantly more complicated (if not impossible) to create in a standard decision analysis framework. Many of the large, complicated models that have been created have required a substantial amount of development work,

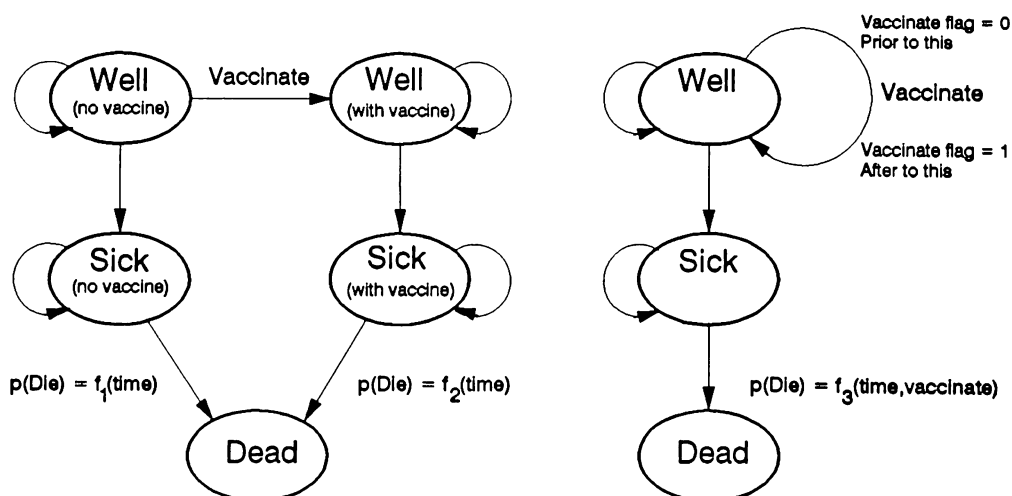


Figure 3: Standard vs Monte-Carlo Markov Process

and their complexity limits the ability of outside reviewers to assess the model's machinery. For example, the coronary heart disease policy model (Weinstein *et al*, 1987) is a state-transition model with over 5000 separate states, and required multiple programmer-years of development time.

There are, of course a large number of problems in health care delivery that fit the standard simulation mold more completely. For example, estimating the optimal staffing of an emergency room yet assuring it could handle its peak loads (Vassilacopolos, 1985); determining the appropriate number and location of ambulances for a city's Emergency Medical System (Liu and Lee, 1988), modelling the physical and personnel characteristics of an outpatient facility (Levy, Watford and Owen, 1989) and several other similar problems lie in the purview of standard discrete event simulation.

3.3 Model Performance and Validity

Because the Monte Carlo models described above are stochastic representations of a larger, basically deterministic Markov process, one can (for simple models) test the validity of the Monte Carlo construction by comparing the results of the Monte Carlo version of a model to the complete standard Markov Process that it simulates. Figure 4 illustrates the relationship between the number of states being simulated and the variance in the estimate of the "true" value (the answer given by the full Markov process) as a function of the number of iterations used to evaluate the model. As shown in the figure, preliminary results indicate that relatively few iterations are required to produce results sufficient to carry out sensitivity analyses, debug models, etc. Furthermore, although the variance in the estimate from the "true" model decreases rapidly as the number of iterations increases, the accuracy of the model does not degrade as the number of states being represented increases. In fact, the early empirical results from our work indicate that the more states being represented, the *smaller* the error for any given number of iterations. It is possible that this occurs because as models grow more complex, there are more chances in any given iteration for the specific member of the cohort to "regress to the mean". Since the final life expectancy for any given member of the cohort is the result of multiple draws from several different underlying distributions, a member of the cohort would have to draw outliers on several different distributions to move very far away from the mean.

Like many discrete event simulations, standard decision analysis models are often difficult to validate. In addition to the use of content experts to assess face

validity (to assure that all possible clinical outcomes have been modelled, that the factors effecting success/failure are well described, etc) the most common method of validating decision analysis models is comparing the answer of the model under a set of conditions for which the answer is either known from empirical work or can be presumed from inspection. For example, in a decision between medical and surgical therapy for a disease, if one therapy is set to be both more effective and less risky, it should dominate the other. Repetitive sensitivity analyses, where various parameters are set to their extremes, (and the implied answer is known) provide confidence that the model behaves as clinically expected.

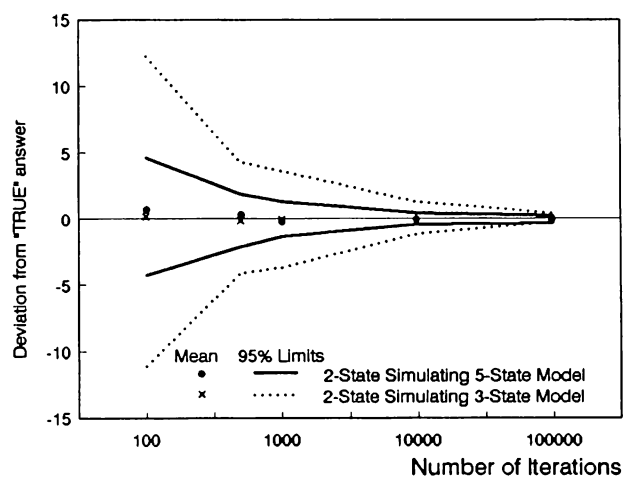


Figure 4: Model performance. Mean and 95% confidence limits on the results of Monte Carlo simulation compared to the "true" model results (baseline) as a function of the number of iterations.

One substantial advantage of the Monte Carlo method is that it allows for a more direct inclusion of continuous risk prediction variables than standard Markov processes. In standard processes, any risk stratification that is incorporated directly into the structure of the model must be accomplished by creating separate states, as in Figure 3. This requires breaking up continuous variables into dichotomous or polychotomous categories, and the relevant parts of the model replicated by a separate set of states.

4 EXAMPLE: END-STAGE LIVER DISEASE (ESLD)

At the Deaconess Hospital, we are developing a simulation model of end-stage liver disease to evaluate the optimal timing of transplantation. The model is an

example of an analysis designed to understand the optimal timing of a technologic intervention in a chronic disease. In general, these "timing" problems are large, complicated, are unlikely to be solved by randomized controlled trials, and may rely on other observational methods to evaluate therapeutic strategies. Therefore, we created a decision analysis to evaluate the transplant/no transplant decision and search over a set of selection criteria to find the optimal timing of transplantation. Previous work at the Deaconess had shown that survival at and after transplantation was dependent upon several different clinical pre-operative factors (see Table 2, from Roberts, 1989). Because the level of these predictors changes over the course of disease, multiple different states were needed to define various possible combinations of pre-operative clinical conditions under which a transplantation might occur. Initial efforts to create a standard Markov process model of the problem resulted in a model with over 2000 states, even after many of the risk prediction strata were reduced from polychotomous to dichotomous categories.

Table 2: Predictors of post-transplant survival by stage

24-Hour	30-Day	Long-Term
●Creatinine ●Prior RUQ surgery	●Creatinine ●Life Support	●Re-transplant ●Crossmatch reaction ●Infection ●Bilirubin ●Malnutrition

4.1 State Based Simulation Model

The basic model is illustrated in Figure 5. The model was created using *SMLTREE*[®], a standard decision analysis package (Hollenberg, 1990). Patients begin the model clinically stable with end-stage liver disease (STABLE WITH ESLD). At the beginning of each cycle, a boolean statement is evaluated to determine if any event in the prior cycle makes the patient a candidate for emergent transplantation. If so; the patient is listed for transplant and enters the transplant queue (AWAIT TRANSPLANT). If the patient fails the emergent transplant selection criteria, a similar boolean statement is evaluated to determine if the patient passes the elective transplant criteria currently being evaluated. If the patient does not pass the elective transplant criteria, control passes to a standard decision tree that

determines whether events that may develop while the disease progresses occur during that cycle. In addition to specific events (generally complications) the model "ages" the value of the clinical predictors that are necessary to calculate future operative risk when and if the patient comes to transplantation, as well as allow for other specific events to occur. The patient may die of natural causes (DEATH).

Once the patient enters the waiting queue, he may receive an organ or have a complication while waiting. Complications are similar to the complication tree attached to the STABLE WITH ESLD state, but specific complications occur at different rates, and are a function of the current value of several clinical variables. The complication may be fatal (DEATH), or may simply change the value of the variables that predict transplant survival. If the patient is transplanted, he may die from the operation (DEATH) or survive (STABLE AFTER TRANSPLANT). The probability of operative death and first 30-day death are calculated as the result of two logistic regression equations which have been estimated on the transplantation experience at the Deaconess Hospital. The actual probability is calculated from the current value of the clinical predictor variables at the time of transplantation. If the patient survives the perioperative period, they then may live out their life post transplant. During this time they may die (DEATH), or have graft failure and either die (DEATH) or require a re-transplantation (AWAIT TRANSPLANTATION).

The evaluation of the model is carried out by thousands of iterations; the result is an average life expectancy for patients proceeding through the model given the specific elective selection criteria specified. As the criteria are changed, the average life expectancy changes; a large set of criteria is searched over, that criteria with the longest life expectancy is deemed the optimal selection criteria.

4.2 Model Calibration

One of the reasons that decision analysis was used to evaluate this problem is that there is no single, complete set of data available to answer the timing question. Current published reports of the effectiveness of liver transplantation have been somewhat more simplistic, comparing the actual survival of patients with transplantation to the *expected* survival of a clinically similar group of patients based on Cox proportional hazards models. The difficulty with this modelling approach is that it only compares one specific selection strategy (that used by the transplant center) to estimate if transplantation increases survival over the natural

history of disease at the time transplantation occurred. It tells us nothing about whether or not the patients who were transplanted could have survived longer if they had been transplanted earlier or later in the course of their disease. It is the ability to predict changes in outcomes based on changes in the selection criteria that gives decision analytic approaches to this problem their potential.

However, modelling the relationship between selection criteria and outcome requires not only the ability to predict survival based on clinical variables, it also requires quantitative knowledge of how those predictor variables change during the natural course of disease. Consequently, data must come from very disparate sources, and be integrated by the structure of the model. The model requires data to calibrate risk prediction equations (logistic regression and Cox proportional hazards models) that create estimates of the survival time with and without transplantation given the set of clinical characteristics that exist at the time of transplantation. Original calibration for the

transplantation survival section came from our own experience at the Deaconess Hospital, but small sample sizes limited the number of clinically relevant variables that could be accurately estimated. Recent acquisition of the United Network Organ Sharing (UNOS) database of over 12,000 transplants with several years of follow-up will allow for more precise calibration of the model.

Unique to this modelling effort is the requirement of a quantitative description of the natural history of end-stage liver disease. Unlike standard textbook descriptions of natural history, the simulation model requires data on how the clinical predictors of transplantation success change over time, either as rates of change or probabilities of a certain change in a given period of time. In addition to literature reviews, the time course of relevant variables will be extracted from CLINQUERY, a large clinical database containing over 8 years of data on over 100,000 general medical patients admitted to the Beth Israel Hospital in Boston, some 3-4000 of whom have chronic liver disease (Shaffran, 1989).

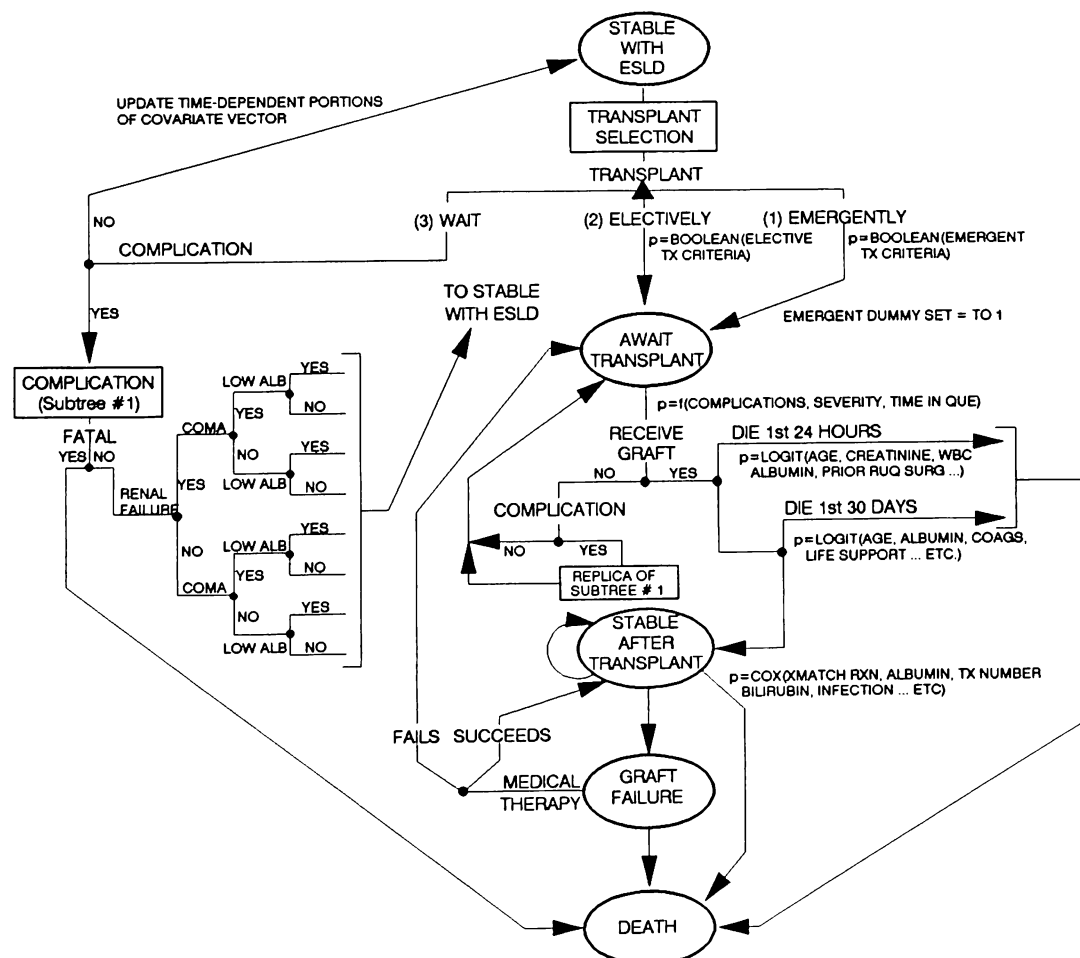


Figure 5: Basic Markov model, end stage liver disease

4.3 Model Caveats

The major caveat to the current version of this model is that it is very coarsely calibrated, and the accuracy of any currently calculated numeric results are suspect. More fundamental, however, is the fact that it is very difficult to test the validity of the model as a whole. When fully calibrated, we expect to attempt limited prospective confirmation of the model on current patients being evaluated for transplantation.

5 CONCLUSIONS

This paper has shown that compact, straightforward Monte Carlo models based on Markov processes can duplicate the results produced by a fully described standard Markov process. There are several advantages to the conversion of Markov Process models into simulation models. First, the structure of the simulation model is substantially simpler than the corresponding Markov process it represents. This makes the basic layout of the model easier to describe and understand (especially to the non-quantitative clinician), and may make the model easier to "debug". Secondly, by releasing the "lack of memory" constraint that limits standard Markov processes, models with much more clinical richness can be practically modelled. The particular methods presented used a standard decision analysis software (SMLTREE) and through characteristics of its Monte Carlo sensitivity analysis converted standard Markov processes into simulation models; hence the presentation and analysis of the model occurs in the context of a software system widely used by medical decision analysts. Finally, and perhaps more importantly, the model allows for the direct inclusion of *continuous data* to make risk or transition probabilities; in standard Markov processes these variables must often be broken down into di- or polychotomous categories; each level of the variable represented by a separate set of states.

The model has yet to produce usable "answers" to the problem of the optimal timing of liver transplantation in end-stage liver disease because of the lack of adequate data to calibrate the model. As work continues and the calibration becomes more precise and based on datasets of sufficient sample size, the model will be used to make predictions concerning specific selection criteria. To date, the main benefit of the modelling system is that it has shown that simulation modelling is an appropriate extension of a standard decision analysis methodology for modeling time-varying events.

REFERENCES

- Beck, J. R. and Pauker, S. G. 1983. The Markov process in medical prognosis. *Medical Decision Making* 3:419-58.
- Dittus, R. S., Roberts, S. D., and Wilson, J. R., 1989. Quantifying uncertainty in medical decisions. *Journal of the American College of Cardiology*. 14(3): 23A-28A.
- Hollenberg, J. P. 1984. Markov cycle trees: a new representation for complex Markov processes. (abstract) *Medical Decision Making* 4:529.
- Hollenberg, J. P. 1990. SMLTREE The all purpose decision tree builder. version 2.9. The Pratt Medical Group, Boston, MA
- Levy, J.L., Watford, B.A., and Owen, V.T. 1989. Simulation analysis of an outpatient services facility. *Journal of the Society for Health Systems*. 1(2):35-49.
- Liu, S.L. and Lee, J.T. 1988. A simulation of a hospital emergency call system using SLAM II. *Simulation*. 51(6):216-221.
- Roberts, M. S. et al 1989. Predictors of intraoperative death and long-term survival in liver transplantation: a multistage analysis. *Transplantation Proceedings* 21(1): 2439-2442.
- Roberts, S.D. and Klein, R. W. 1984. The simulation of logical networks (SLN). *Simulation* 43:224-232.
- Shafran, C. 1989. ClinQuery: A system for online searching of data in a teaching hospital. *Annals of Internal Medicine* 111(9):751-756.
- Vassiliacopolos G. 1985. Allocating doctors to shifts in an accident and emergency department. *Journal of the Operational Research Society*. 36(6):517-523
- Weinstein, M.C., et al 1987. Forecasting coronary heart disease incidence, mortality and cost: The coronary heart disease policy model. *American J of Public Health* 77:1417-1426.

AUTHOR BIOGRAPHY

MARK S. ROBERTS is an academic general internist and Director of the Center for Clinical Information Systems at the New England Deaconess Hospital and Instructor of Medicine at Harvard Medical School. He received an A.B. in Economics from Harvard College, an M.D. from Tufts Medical School, and M.P.P. (Public Policy) from the Kennedy School of Government at Harvard University. His research interests include decision analysis and the application of quantitative methods to problems in medical care.

Work: (617) 735-0720

Fax: (617) 738-6919

Email: MARKR@HARVARD.HARVARD.EDU