Stochastic Tree Models In Medical Decision Making

GORDON B. HAZEN — *Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, Illinois, 60208-3119 (hazen@iems.nwu.edu)*

JAMES M. PELLISSIER — *Clinical and Health Economic Statistics, Merck Research Laboratories, 10 Sentry Parkway, BL3-2, Blue Bell, Pennsylvania, 19422*

JAYAVEL SOUNDERPANDIAN — *Department of Business, University of Wisconsin-Parkside, Box No. 2000, Kenosha, Wisconsin, 53141-2000*

August 29, 1997

Abstract: The *stochastic tree* is a recently introduced generalization of the decision tree which allows the explicit depiction of temporal uncertainty, while still employing the familiar rollback procedure for decision trees. We offer in this paper an introduction to stochastic tree modeling and techniques involved in their application to medical treatment decisions. We also describe an application of these tools to the analysis of the decision to undergo a total hip replacement from the perspectives of an individual patient (via utility analysis) and of society (via cost-effectiveness analysis).

HEALTH CARE – TREATMENT

DECISION ANALYSIS – APPLICATIONS

UTILITY/PREFERENCE – APPLICATIONS

Should a patient elect to undergo surgery or to be managed medically? Should a managed care company put a drug on its list of approved medications? What medical programs or procedures should governmental agencies reimburse? Individuals and institutions face medical treatment decision problems from a variety of perspectives. Such problems can involve serious health consequences that may extend many years into a patient's future. What measures are most germane to help make these decisions depend on the perspective and time horizon in question. For example, the decision an individual patient faces when covered by insurance is driven by the medical treatment that best accommodates the patient's attitudes about risk and quality of life over the rest of his or her lifetime. The decisions a managed care company confronts are driven by costs over a finite (less than three years) time horizon. Government decisions at the societal level need to incorporate both the costs and the effectiveness of the medical treatment or program over the lifetime of the patient. A number of approaches and techniques exist for assessing the effectiveness of interventions, medications and programs. Randomized controlled clinical trials and epidemiologic studies involve the collection and analysis of primary data. Meta-analytic approaches synthesize information from multiple studies. However, the time horizons appropriate for medical decisions frequently extend beyond the available data. When this is the case, model-based analyses are appropriate. In its recent recommendations, the Panel on Cost-Effectiveness in Health and Medicine convened by the U.S. Public Health Service explicitly states, "Where direct primary or secondary empirical evaluation of effectiveness is not possible, the use of modeling to estimate effectiveness is a valid mode of scientific inquiry" [Gold et al., p. 168].

Researchers have performed a number of stochastic analyses of medical decisions (for example, Mooney [1990]; Pauker [1976]; Ransohoff et al. [1983]; Roach et al. [1988]; Sonnenberg [1985]). The type of stochastic model nearly all have used in these analyses is the discrete-time Markov chain, in which time is treated as a sequence of successive epochs (months, days, or years) during which the state of the system remains fixed. The state may change only at the end of an epoch. By contrast, a continuous-time Markov chain treats time as a continuum, and jumps from one state to another may occur at any instant (see, for example, Cinlar [1982], Ross [1970]). While continuous-time Markov chains lend themselves well conceptually to modeling medical treatment decisions, they have not been used extensively.

For medical analyses, the most commonly used method for the formulation and display of discrete-time Markov chains is the Markov cycle tree developed by Hollenberg [1984]. Hazen [1992] introduced a continuous-time analog of the Markov-cycle tree, the *stochastic tree*, which combines features of decision trees [Raiffa 1968] and stochastic-process transition diagrams. Stochastic tree diagrams not only can depict continuously distributed temporal uncertainties, but, like decision trees, can be rolled back to determine optimal decisions. This rollback procedure is an instance of the method of successive approximations from stochastic dynamic programming (Denardo [1982], Ross [1970]). It can be applied not only for risk-neutral objective functions, such as mean quality adjusted lifetime [Hazen 1992], but also for the calculation of expected utility [Hazen and Pellissier [1996].

One of the first applications of stochastic tree modeling in medical decision making was an investigation of total hip replacement surgery for patients with severe osteoarthritis

(OA) of the hip. An estimated 120,000 hip replacements are performed per year in North America [Harris and Sledge 1990]. This decision is interesting from a number of points of view. Although total hip replacement surgery (total hip arthroplasty, THA) is the most successful procedure for patients with advanced arthritis of the hip, the treatment decision for an individual patient with this condition can be difficult. While the surgery offers the potential of increased quality of life, it carries additional risk of mortality and morbidity. From a societal perspective, given the increasing incidence of severe osteoarthritis of the hip, the growing demand for THA, and the high costs associated with the procedure, there is concern that a larger share of health care resources will be spent on THA in the future. Rising costs and attempts to improve the efficiency of health care delivery have stimulated the need for analyses that objectively examine the trade-offs between therapeutic benefits and expenditures. If the economic consequences of surgical interventions are included, procedures such as THA are particularly vulnerable to scrutiny for several reasons. First, it is an elective procedure. Second, its target population is largely geriatric. Third, it is an expensive, high-technology procedure. Finally, in terms of short-term costs alone, medical management would appear to represent a less expensive treatment alternative for OA of the hip. This raises the following question from a societal as well as a policy-making perspective: "How cost-effective is THA compared to continued medical management?"

Our purpose in this paper is to introduce stochastic trees and discuss techniques (factoring, Coxian background mortality, risk-sensitive preferences, continuous-risk utility assessment) involved in their application to medical treatment decisions. We will illustrate how these tools have been applied to analyze the decision to undergo a total hip

replacement from the perspectives of an individual patient (via utility analysis) and of society (via cost-effectiveness analysis).

Stochastic Trees

In its simplest and most useful form, a stochastic tree is a transition diagram for a continuous-time Markov chain, unfolded into a tree structure. Researchers have used stochastic trees as modeling tools to analyze medical treatment decisions [Chang, Pellissier & Hazen 1996; Gottlob et al. 1995; Hazen 1992; Hazen 1993].

Figure 1 shows a simplified stochastic tree model of nonsurgical treatment of transient ischemic attacks, motivated by Matchar and Pauker [1986]. Wavy arrows are labeled with rates, and signify transitions that take time to accomplish. We call nodes from which wavy arrows emanate *stochastic* nodes. Straight arrows are labeled with probabilities, and signify immediate transitions. Nodes from which straight arrows emanate play the same role as chance nodes in a decision tree. Hazen (1992) presents a complete stochastic tree representation of the Matchar and Pauker model. Hazen and Pellissier (1993) discuss equivalent representations for stochastic trees and methods to simplify their structure.

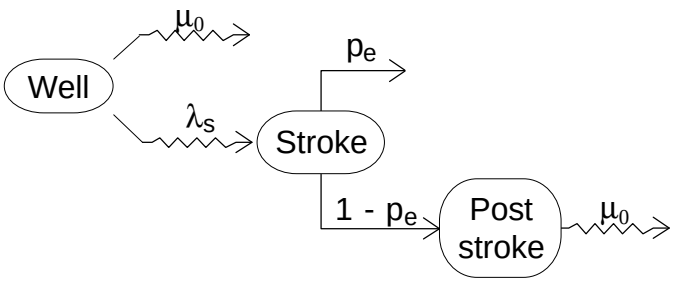

Figure 1: Stochastic tree model for nonsurgical treatment of transient ischemic attacks. The nodes represent states (Well, Stroke, Post Stroke), and the arrows represent transitions between states. The Dead state does not explicitly appear, being modeled by termination of the process.

One popular measure of treatment efficacy used in the medical decision making literature is *mean quality-adjusted duration* [Weinstein et al. 1980]. This measure is calculated by weighting the average duration spent in each particular health state by a *quality rate* proportional to the desirability of that state. Typically the Well state is assigned a quality factor of 1, the dead state 0, and other states intermediate values. The result for each treatment is a duration which, if spent entirely in the Well state, would be equivalent to receiving the treatment and all its consequences. Mean quality-adjusted duration has seen widespread application. For example, Weinstein and Stason [1976] model hypertension treatment; Beck and Pauker [1981] examine anticoagulation therapy; Hillner, Hollenberg and Pauker [1986] investigate estrogen for the prevention of osteoporosis; Plante, Piccirillo and Sofferman [1987] model cancer treatment; Mooney, Mushlin and Phelps [1990] study magnetic resonance imaging for multiple sclerosis.

A further advantage of the stochastic tree model is that it allows the recursive computation of mean quality-adjusted duration by *rolling back* the stochastic tree [Hazen

1992], much as one would roll back a decision tree. We describe this process in an appendix.

Factored Trees

Quite often, a complex medical situation can be thought of as several health processes unfolding in parallel. For example, the stochastic tree of Figure 1 can be viewed as two simultaneous processes, one consisting of background mortality and the other, stroke and its consequences. We can depict these two processes as separate stochastic trees, as is shown in Figure 2: We say we have *factored* the original tree into a background mortality factor and a stroke factor. We can recover the original stochastic tree by forming the Cartesian product of the states in the background mortality factor with the states in the stroke factor, with each product state inheriting all transitions out of the factor states that compose it.

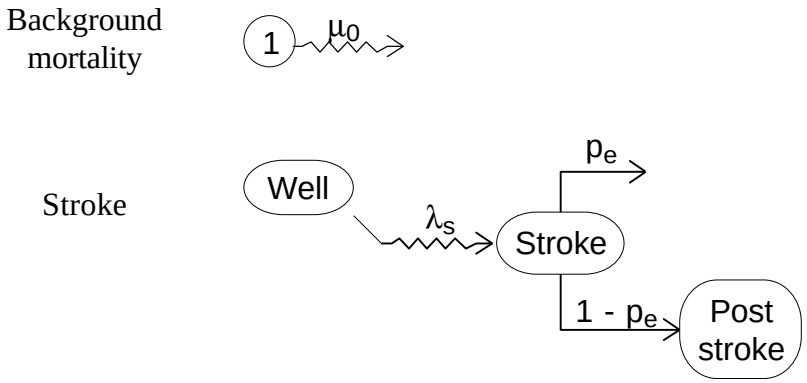

Figure 2: The stochastic tree model of Figure 1 contains a background mortality process and in parallel, a stroke process. In the factored stochastic tree shown here, we depict these processes separately.

This factoring approach facilitates modeling by allowing the modeler to focus separately on different problem aspects and refine or modify them as required. For example, the assumption of exponentially distributed human lifetime depicted in the simple background mortality factor above is usually unacceptable. A background mortality factor that more accurately depicts human lifetime distributions is the Coxian model [Cox 1955] shown in Figure 3. The combination of this Coxian mortality factor with the original stroke factor is an improved model of disabling stroke.

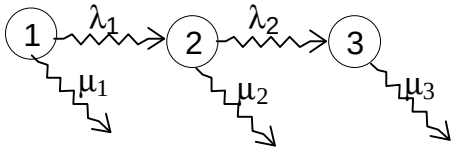

Figure 3: The Coxian model offers a more detailed look at human lifetime. Here the individual may be thought of as proceeding through three life stages, each with a different mortality rate. The parameters and the number of stages can be chosen so as to very accurately approximate human survival distributions.

In this model, the risk of stroke is assumed to be independent of the background risk of mortality. In general, this needn't be the case, that is, the stroke likelihood may be a function of the Coxian stage. For example, life-threatening risks, such as diet or smoking, also affect the stroke risk. Such a situation might be modeled by letting the stroke rate λ_s be higher for later, less healthy Coxian stages.

It is much easier to construct the two factors separately rather than to attempt to directly construct their Cartesian product. For the stroke model, this Cartesian product is shown in Figure 4. It is the product tree on which rollback operations must occur, but the

modeler can employ implicit rollback methods [Hazen 1992, Hazen 1993] to achieve rollback without ever graphically formulating the product tree. Stochastic tree software is now available [Hazen 1997] for formulating and solving factored stochastic trees, and may be obtained by request from the authors.

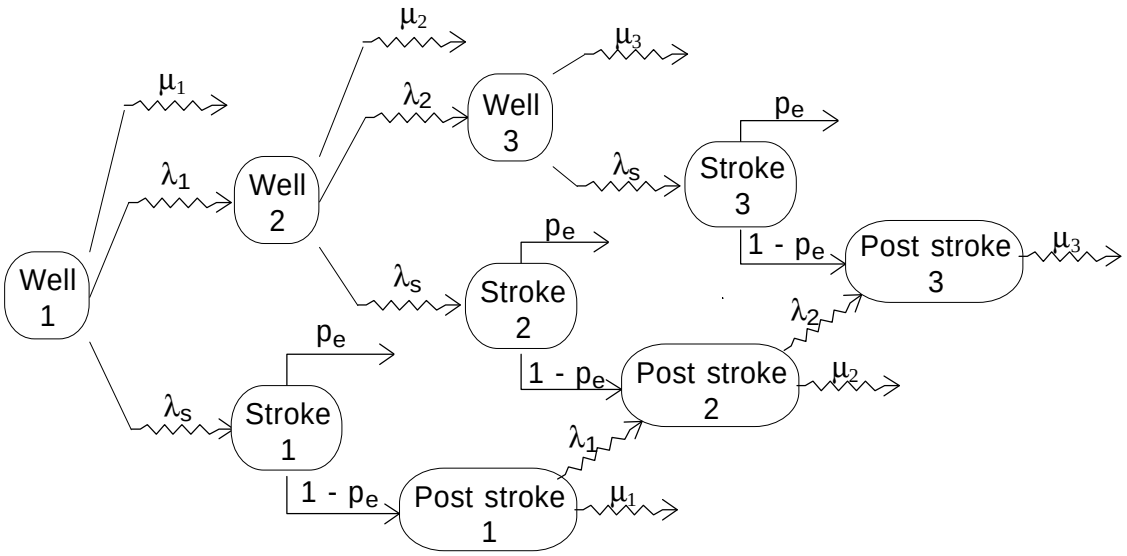

Figure 4: The state descriptors of the product tree indicate the stroke factor state and the stage of the Coxian mortality factor. For example, a patient in Post Stroke 2 is in the second Coxian life stage and the Post Stroke state. The patient can enter Post Stroke 2 by being in the Post Stroke 1 state and having the mortality factor advance to stage 2, or by surviving a stroke that occurs in stage 2.

The factor approach facilitates model formulation by decreasing the chance of modeling errors and improves model presentation by allowing a simpler and more easily grasped display. Multifactor trees arise naturally. For example, Hazen [1993] formulates the Tsevat *et al.* [1986] model of warfarin treatment for cardiomyopathy into a stochastic

tree having factors representing background mortality, systemic embolism, pulmonary embolism, systemic hemorrhage, and anticoagulant status.

Utility Functions Over Stochastic Trees

Stochastic tree rollback is possible for risk-sensitive as well as for risk-neutral preferences. Hazen and Pellissier (1996) present von Neumann-Morgenstern utility functions over stochastic trees that allow convenient rollback. One such utility function is a risk-sensitive extension of quality-adjusted duration that allows distinct constant coefficients of risk aversion across health states. We discuss this utility function and a method for rollback in an appendix.

Continuous-Risk Utility Assessment

Continuous-risk utility assessment protocols differ from standard preference assessment protocols by offering choices between alternatives in which risks are present in continuous time. For medical applications, continuous-risk utility assessment scenarios are both more realistic and more familiar than other simpler but more artificial standard gamble approaches.

To understand how continuous-risk utility assessment works, consider the scenario in Figure 5, used to assess a subject's risk attitude about future time spent in the American College of Rheumatology (ACR) functional class IV (incapacitated) for osteoarthritis of the hip. Here the subject is asked to specify what *compensating immediate mortality risk* p she is willing to incur to decrease her mortality rate from the high rate μ_1 to a lower rate μ_0. The subject's response determines her coefficient of risk aversion for health state ACR IV.

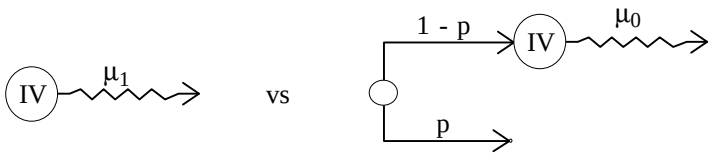

Figure 5: In this continuous-risk assessment scenario, a patient must specify what chance p at immediate death he is willing to tolerate to improve her subsequent mortality rate from μ_1 to μ_0.

Medical personnel present continuous-risk assessment scenarios to patients using graphs of survival curves. For example, when $\mu_0 = 0.04386$/yr, $\mu_1 = 2\mu_0 = 0.08772$/yr, and p = .18, the choice is between the survival curves S_0 and S_1 in Figure 6. Indifference would indicate the subject is willing in this case to take an 18% chance at immediate death in order to cut her ongoing mortality rate in half. The value of p that should be chosen if the patient is concerned with only average lifetime is $p = p_0 = 0.50$. The indifference probability (p=0.18) for this hypothetical subject is less than the risk neutral probability (p=0.50), indicating that she is risk averse.

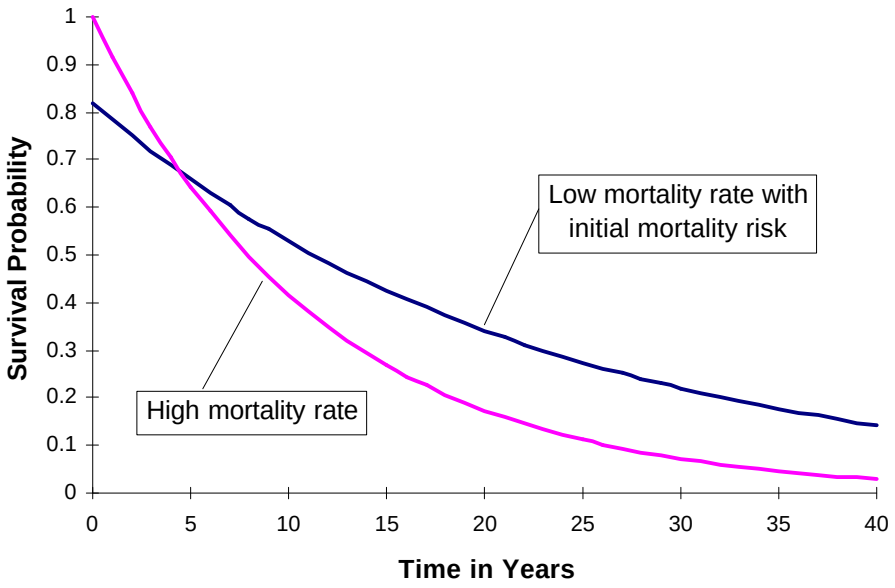

Figure 6: Medical personnel present a continuous-risk assessment scenario to patients using a survival curve representation such as this. By manipulating the initial mortality risk, a patient can move the dashed curve vertically until she is indifferent between the two curves.

Hazen, Hopp and Pellissier [1991] and Pellissier and Hazen [1994] performed continuous-risk assessments using an interactive computer program written in Pascal. It featured graphical and tabular displays of the survival curves corresponding to both options. Subjects could interactively manipulate the probability p of immediate mortality and view the corresponding changes in survival curve, until they reached an indifference value of p. Subjects do successfully respond to the seemingly more challenging continuous-risk questions and in fact react favorably to the aspects of realism and familiarity present in such scenarios [Pellissier and Hazen 1994].

Risk Attitude and Individual Treatment Choice

We illustrate continuous-risk utility assessment and stochastic tree rollback through the following (simplified) hypothetical version of the total hip replacement decision faced by an individual patient. A 60-year-old woman at the time of decision occupies functional class III (limited capabilities). Without surgery, she will progress to class IV (disabled) at rate $\lambda = 0.0249$/yr (Figure 7). With surgery, she will be returned permanently to functional class I (healthy), unless there are fatal surgical complications. These occur with probability 25%. Her age and gender specific mortality rate due to other causes is $\mu_0 = 0.04386$/yr.

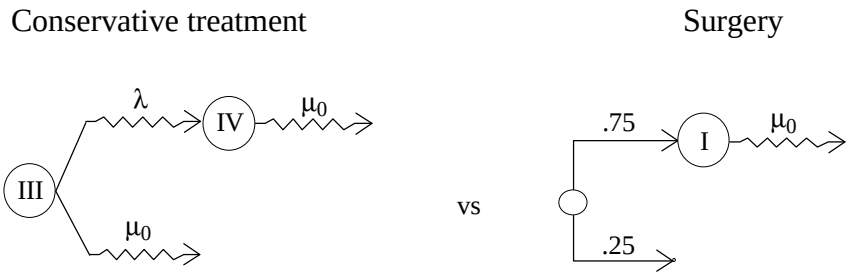

Figure 7: A hypothetical choice between conservative treatment and surgery is depicted by a pair of stochastic trees.

Let the *lifetime valuation* of a health state be the expected utility of a lifetime occupancy of that state, assuming mortality rate μ_0. We can assess lifetime valuations using continuous risk utility assessment. We can arbitrarily set the lifetime valuation of ACR functional class I to 1.0. Suppose this woman assigns lifetime valuations 0.92 and 0.30 to the ACR functional classes III and IV. Should the woman be risk neutral in all states, then

rollback formulas can be applied to yield overall expected utilities (equal to mean quality-adjusted durations, due to risk neutrality). The results are

$U_{conserv}$ = 15.857 yr.

U_{surg} = 17.100 yr.

giving a year-and-a-quarter advantage to the surgical option. On the other hand, should the woman be risk averse with identical coefficients of risk aversion 0.156 for all states, then rollback gives

$U_{conserv}$ = .8513 $CE_{conserv}$ = 7.002 yr.

U_{surg} = .7500 CE_{surg} = 5.645 yr.

Here $CE_{conserv}$ and CE_{surg} are *certainty equivalent lifetimes* in the healthy class I. For example, the utility of 7.002 years spent in class I is equal to 0.8513, the expected utility of the conservative strategy. Therefore the certainty equivalent lifetime of the conservative strategy is $CE_{conserv}$ = 7.002 years. This is a 1.4 year advantage over the 5.645 year certainty equivalent lifetime of the surgical strategy. Under risk neutrality, certainty-equivalent lifetime and mean quality-adjusted duration coincide. In this hypothetical example, the change from risk neutrality to risk aversion results in a switch from the more risky surgical strategy to the safer conservative strategy.

The coefficients of risk aversion 0.156 for health states I,III,IV correspond to a compensating immediate mortality risk p = 0.18 in the continuous-risk assessment scenario of Figures 5 and 6. The effect of risk attitude on treatment decision can be captured by graphing the certainty-equivalent lifetimes of the competing strategies against the assessed compensating immediate mortality risk p, as is done in Figure 8. A value of p = 0.40 would result in a tossup between the conservative and surgical treatments.

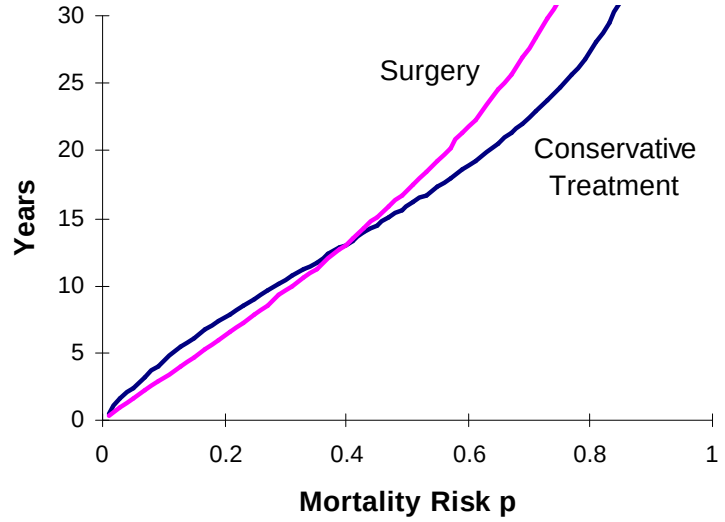

Figure 8: A graph of certainty-equivalent lifetime (years) versus assessed compensating immediate mortality risk p reveals that the hypothetical patient prefers surgery if she selected a compensating immediate mortality risk less than 0.40, and prefers conservative treatment if her compensating risk exceeds 0.40. A risk-neutral individual (one concerned only with average lifetime) would have compensating immediate mortality risk 0.50.

The Total Hip Replacement Decision

We have developed stochastic tree models for the hip replacement decision problem. These models were constructed in collaboration with Rowland Chang of Northwestern University Hospital (Chang, Pellissier and Hazen 1996).

We chose functional outcome as the primary measure of effectiveness for total hip arthroplasty (THA). For modeling purposes, we adapted the four-state American College of Rheumatology (ACR) functional status classification for use in OA of the hip [Steinbroker *et al.* 1949]:

CLASS I: Complete ability to carry on all usual duties without handicaps.

CLASS II: Adequate for normal activities despite handicap of discomfort or limited motion in one or more joints.

CLASS III: Limited only to little or none of duties of usual occupation or self care.

CLASS IV: Incapacitated, largely or wholly bedridden or confined to wheelchair; little or no self care.

A candidate for a total hip replacement is typically in functional class III.

Our stochastic tree models incorporate the best surgical probability and failure rate information available from the medical literature. When published data are highly disparate, we used the values that most biased *against* surgery. Chang, Pellissier and Hazen [1996] discuss the data and assumptions that underlie this model.

Conservative Treatment versus Hip Replacement Surgery

The alternative to surgery on an arthritic hip is conservative management, which seeks to reduce the pain and discomfort associated with deterioration in the patient's functional abilities. Because osteoarthritis of the hip is a degenerative disease, such deterioration is inevitable and can only be interrupted by death. Figure 9 is a stochastic tree factor depicting the consequences of the conservative treatment. From results reported by van Saase [1989], we estimated the constant yearly rate λ of natural progression of OA of the hip from ACR functional class III to class IV to be $\lambda = 0.0249$. To complete the model, we include a Coxian background mortality factor. The full tree is the Cartesian product of these two factors.

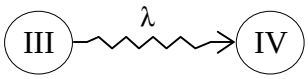

Figure 9: In our conservative treatment factor for osteoarthritis of the hip, a patient moves from functional class III to functional class IV at a rate of λ = 0.0249 occurrences per year.

In THA, the surgeon removes the arthritic joint and replaces the joint with a prosthesis. Although THA entails far more uncertainty than conservative management, the potential rewards in terms of restored functional capabilities are great. A successful THA will nearly restore a patient's full function. However, an unsuccessful operation can result in such dire consequences as insidious infection or death.

Uncertainty exists even when an operation has been successful. The presence of infection might cause the prosthesis to fail over time (*septic* failure). The prosthesis might fail for a variety of other reasons such as mechanical loosening, prosthesis breakage or dislocation (*aseptic* failure). In the event of any of these failures, a revision surgery is necessary.

The risks involved in revision surgeries are typically greater than those in primary THA surgery, and successful outcomes are less likely. Also, a successful revision surgery will not result in the same magnitude of functional well-being as a successful primary surgery. Given the results of a revision, the same process of joint and prosthesis deterioration can again force a need for further revision. The outcome probabilities in revision surgeries are dependent on a patient's previous history. Particularly important is whether a patient has ever experienced an infection. At some point, a patient's existing bone stock and the condition of his or her hip does not allow for further revision. In our

model, we assumed no patient would receive more than three revisions. Figure 10 shows the surgery factor, depicting primary THA, possible aseptic and infection revision surgeries, and associated outcomes. The full surgery model is the Cartesian product of the surgery factor with a Coxian background mortality factor. Table 1 summarizes the outcome probabilities associated with primary surgery, aseptic revisions and infection revisions. Table 2 shows yearly revision rates following surgeries.

		ACR Class Outcomes				
Surgery	Initial ACR Class	I	II	III	IV	Surgical Mortality
Initial THA[a]	III	0.6925	0.2430	0.0600	0	0.0050
Aseptic						
Revision 1[b] (no septic history)	III	0	0.7015	0.2865	0	0.0120
Revision 1[b] (septic history)	IV	0	0	0.7015	0.2865	0.0120
Revision 2[c] (no septic history)	III	0	0.6363	0.3517	0	0.0120
Revision 2[c] (septic history	IV	0	0	0.6363	0.3517	0.0120
Revision 3[c] (no septic history)	III	0	0.8477	0.1403	0	0.0120
Infection Revision[d]	IV	0	0	0.7692	0.2115	0.0193

[a] Derived from Kavanagh et al. (1989) and the *Orthopedic Knowledge Update IV* (1992)
[b] Derived from Kavanagh and Fitzgerald (1985)
[c]Derived from Kavanagh and Fitzdgerald (1987)
[d] Sanzen et al. (1988)

Table 1: In total hip arthroplasty (THA), functional class following surgery depends on current functional class and the number of previous surgeries.

.

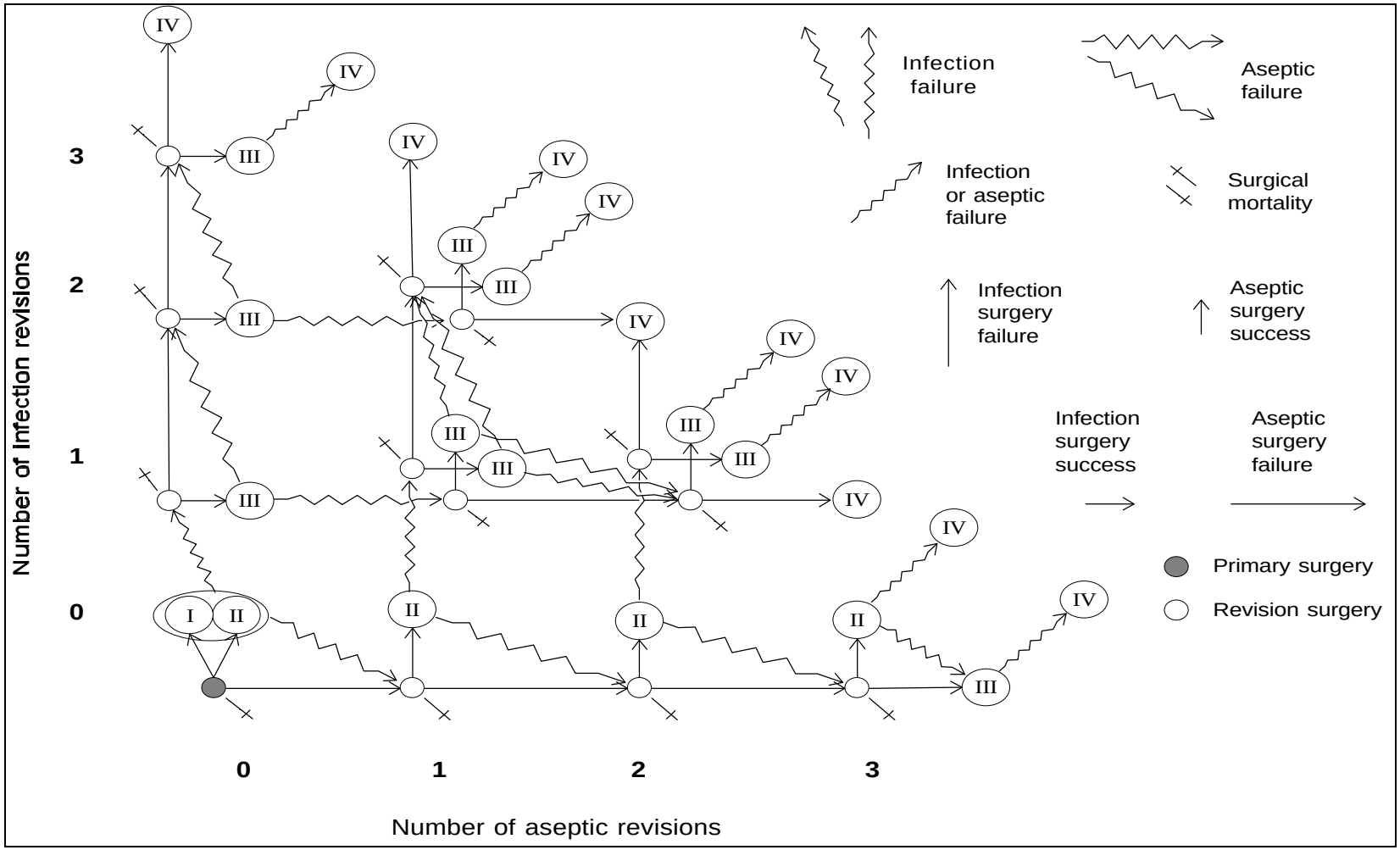

Figure 10: This stochastic factor describes patient prognosis following primary THA surgery. A patient begins at the primary surgery node in the lower left corner and over time will move rightward and upward until death occurs or functional class IV is reached.

Surgery	Aseptic Failure Rate	Infection Failure Rate
Initial THA	0.01	0.002
Aseptic revision 1	0.04	0.02
Aseptic revision 2	0.05	0.035
Aseptic revision 3	0.10	0.05
Infection revision 1	0.04	0.02
Infection revision 2	0.05	0.035
Infection revision 3	0.10	0.05

Table 2: Annual prosthesis failure rates depend on the identity (aseptic or infection) and number (initial, first, second, third) of the most recent surgery. Data was obtained from Chang, Pellissier and Hazen (1996).

Analysis of Individual THA Decisions

Pellissier and Hazen [1994] elicited preferences and risk attitudes over functional outcome. The subjects in their study exhibited a wide range of preferences and risk attitudes: Risk-seeking behavior, various levels of risk aversion and disparate preferences over functional status. Based on expected utility analysis, 44 of 45 subjects preferred the total hip replacement surgery. This preference proved to be quite robust over variations in value trade-offs, risk attitudes, and model parameters. Responses and model parameters would need to change by orders of magnitude to reverse this treatment recommendation [Pellissier et al., 1996].

Here, for example, is a representative case in which a risk-neutral 60 year-old white woman with advanced osteoarthritis of the hip is considering a total hip replacement. Her preferences over the ACR classes of functional well-being are

$$v(I) = 1 \qquad v(II) = 0.8 \qquad v(III) = 0.5 \qquad v(IV) = 0.3.$$

The results of expected utility analysis suggest that this patient should prefer surgery to no surgery. The THA alternative provides 19.84 mean quality adjusted life years while the No THA alternative provides 9.76. The expected times spent in each of the functional classes for both the THA and No THA treatment options appear in Table 3.

	Class I	**Class II**	**Class III**	**Class IV**
No THA	0	0	14.88	7.73
THA	13.27	7.13	1.38	0.60

Table 3: For a 60 year old white female initially in functional class III, the expected time spent in each functional class shows a clear benefit for surgery (THA).

If we assume her risk attitude is the same over all health states, then given this woman's valuations of the functional classes, no change in risk attitude would switch the treatment recommendation from surgery to nonsurgery. Figure 11 shows the relationship between compensating immediate mortality risk p as a measure of risk attitude and the certainty equivalent lifetimes of THA and conservative treatment. Similarly, the surgical recommendation remains optimal for virtually all quality rates satisfying $v(I) \geq v(II) \geq v(III) \geq v(IV)$.

The surgical recommendation for this woman is also insensitive to variation of the rates and probabilities in the model. For example, the rate of deterioration from class III

to class IV without surgery could drop from its base level 0.0249 per year to zero without affecting the decision. An increase of over 65 *times* the surgical mortality risk is required for the treatment recommendation to change. Increases in all prosthesis failure rates (due to both aseptic and septic causes) by a factor of 9.59 are required before the surgical recommendation becomes suboptimal.

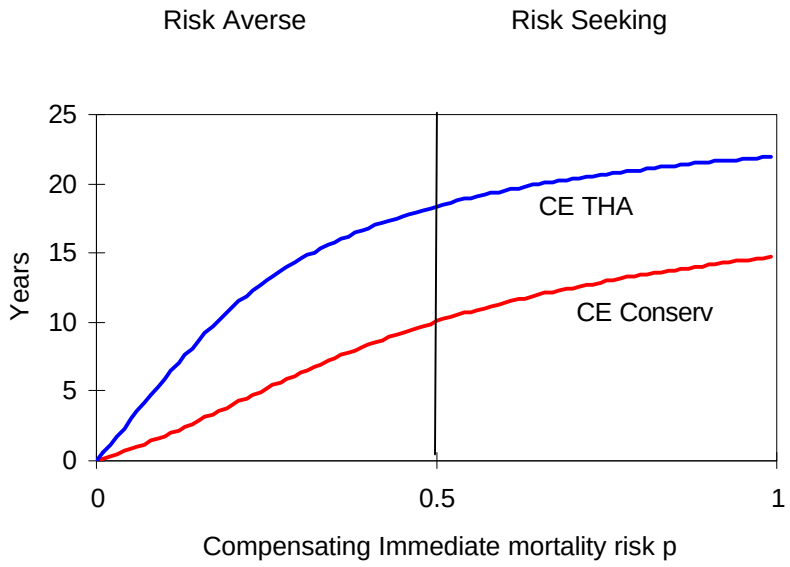

Figure 11: A graph of certainty equivalent (CE) lifetimes versus compensating immediate mortality risk indicates that a patient should prefer hip surgery (THA) over conservative therapy (Conserv) regardless of risk attitude.

Cost-Effectiveness Analysis

Chang, Pellissier and Hazen [1996] have used the THA stochastic-tree models to examine the cost-effectiveness of THA surgery. They perform cost rollback of the trees is performed by letting the parameter v(x) assume the role of cost rate (such as for

medications and home care) assigned to functional class x and treating surgical and immediate rehabilitation costs as tolls. We obtained cost figures from Northwestern Memorial Hospital data (Table 4).

	No THA	Primary THA	Revision
Hospital cost per case	$0	$17,000	$20,000
Physician reimbursement per case	$0	$5,000	$5,000
Rehabilitation cost per case	$0	$3,000	$3,000
Medical costs per year for ACR class III and IV	$775	$775	$775
Custodial care costs per year for ACR class IV	$35,000		

Table 4: The costs associated with the total hip arthroplasty (THA) decision were obtained from Northwestern Memorial Hospital

We did separate analyses were done for women and men in each of four age categories: 60, 70, 80, and 85 and older. For each of these eight categories, we considered a base case using the best-estimate values for probabilities, rates, utilities, costs and discount rate. In addition, we did worst-case analyses to assess the sensitivity of the cost-effectiveness ratios given the uncertainty of the base-case probability and rate estimates. They provide a conservative boundary to the estimates of the age-and gender-specific THA cost-effectiveness ratios.

Base-case cost-effectiveness analysis results for a 60-year-old white female indicate that THA is *cost saving* over her expected lifetime. With both costs and utilities discounted at 3%, THA provided 13.7 QALYs versus 6.82 QALYs for the nonoperative

strategy. The expected lifetime costs of the operative and the nonoperative strategies were $47,649 and $165,440 respectively (in 1991 US dollars). THA is therefore cost-beneficial — it both increases quality and reduces cost. This result is due to the high immediate cost of THA surgery being offset by higher expected costs from custodial care (in class IV) if the arthritis progresses. We found the highest cost-effectiveness ratios for men aged 85 and older. For this group, THA provided 4.16 QALYs versus 2.16 QALYs for the nonoperative strategy. The expected lifetime costs of the operative and the nonoperative strategies respectively were $30,580 and $21,432 resulting in a marginal cost per QALY gained of $4,575. The worst-case model for 60 year-old white women projects that THA is not cost-saving and that the marginal cost-effectiveness ratio is $27,000 per QALY gained. For white men aged 85 years or more, the worst-case cost-effectiveness ratio is nearly $80,000 per QALY gained.

The cost-effectiveness of THA is similar to or better than those of two widely accepted and costly technologies that extend life: coronary artery bypass surgery ($8,100 per quality-adjusted year of life to treat symptomatic 2 vessel disease [Weinstein and Stason 1982]) and renal dialysis ($59,400 to $68,300 per year of life, unadjusted for quality [Garner and Dardis 1987])(Table 5). Thus, in any policy decision in which cost-effectiveness is the criterion, THA should be considered as effective or more effective than these procedures as a means of adding to one's quality adjusted life expectancy. It is noteworthy that THA surgery adds to quality-adjusted life expectancy by improving quality rather than by extending life.

Conclusion

The stochastic tree is a convenient graphical summary of decision problems that have consequences incurred over time, as is the case with most medical treatment decisions. Like the decision tree, it admits a recursive solution algorithm which is suggested by the tree structure. The construction of stochastic trees in factored form considerably simplifies the tasks of model formulation and presentation. The use of Coxian background mortality factors allows for the accurate and convenient representation of human mortality. Risk-sensitive preferences can also be treated in a way that retains the convenient rollback algorithm of the risk-neutral case. Continuous-risk utility assessment arises naturally in the stochastic tree context and allows assessment protocols that are both more realistic and more familiar to subjects than standard gamble techniques.

Stochastic tree modeling has been used for both utility and cost-effectiveness analyses of complex medical treatment decision problems such as the decision to undergo total hip replacement presented here. Gottlob et al. [1995] used stochastic tree methodologies to analyze the cost-effectiveness of total knee replacement surgery. Gottlob and other researchers at the Hughston Sports Medicine Clinic in Atlanta and the Illinois Bone and Joint Institute are also in the final stages of a cost-effectiveness analyses of anterior cruciate ligament reconstruction surgery using similar methodology. Projects relying on continuous-risk utility assessment and stochastic-tree methodologies to aid individual decision making are also ongoing. These include protocols for the choice between mastectomy and lumpectomy for patients diagnosed with ductile carcinoma in

situ, and investigations of differences in risk attitudes and values of patients, rheumatologists and orthopedic surgeons regarding total hip arthroplasty.

[a]Procedure	[b]Additional Cost, $
THA—lifetime estimate (Chang, Pellissier and Hazen 1996)	
60-year-old white woman	Cost saving
≥85-year-old white man	4575/QALY
Low-dose zidovudine therapy for asymptomatic HIV infection—continuous effect (Schulman et al. 1991)	7800/LY
Coronary artery bypass, left main disease plus angina (Weinstein and Stason 1982)	8100/QALY
THA—first 3 years following surgery (Laupacis et al. 1994)	8700/QALY
Hydrochlorothiazide for hypertension (Edelson et al. 1990)	24,900/LY
Screening mammography, women ≥ 50 y (Mushlin and Fintor 1992)	20,000—50,000/LY
Coronary artery bypass, two-vessel disease plus angina (Weinstein and Stason 1982)	37,400/QALY
Renal dialysis, in-center benefit, men (Garner and Dardis 1987)	59,400—68,300/LY
Low-dose zidovudine therapy for asymptomatic HIV—one time effect (Schulman et al. 1991)	83,600/LY
Cholestyramine for high cholesterol (Kinosian and Eisenberg 1988)	91,200/LY
Captopril for hypertension (Edelson et al. 1990)	98,100/LY
Autologous blood donations for elective THA (Etchason et al. 1995)	218,800/QALY
Screening mammography, women < 50 y (Eddy et al. 1988)	220,400/LY

[a]THA indicates total hip arthroplasty; QALY, quality-adjusted life year; LY, life year; HIV, human immunodeficiency virus

[b]Data as reported in the medical literature adjusted to 1991 US dollars using the medical component of the consumer price index (*Economic Indicators* 1994).

Table 5: Cost-effectiveness ratio estimates of THA and other procedures reveal that THA is one of the most cost-effective of medical procedures when increases in quality of life are included.

References

Beck J.R. and Pauker S.G., 1981, "Anticoagulation and atrial fibrillation in the brachycardia-tachycardia syndrome," *Medical. Decision. Making* Vol. 1, No. 3, pp. 285-301.

Chang R.W., Pellissier J.M., and Hazen G.B., 1996, "A cost-effective analysis of total hip arthroplasty for osteoarthritis of the hip," *Journal of the American Medical Association,* Vol. 275, No. 11, pp. 858-865.

Cox, D.R., 1955, "A use of complex probabilities in the theory of stochastic processes," *Proceedings of the Cambridge Philosophical Society,* Vol. 51, pp. 313-319.

Denardo E.V., 1982, *Dynamic programming: Models and applications*, Prentice Hall, Englewood Cliffs, New Jersey.

Economic Indicators, 1994, US Government Printing Office, Washington DC.

Edelson J.T., Weinstein M.C., Tosteson A.N., Williams L., Lee T.H., Goldman L., 1990, "Long-term cost-effectiveness of various initial monotherapies for mild to moderate hypertension," *Journal of the American Medical Association,* Vol. 263, No. 3, pp. 407-413.

Eddy D.M., Hasselblad V., McGivney W., Hendee W., 1988, "The value of mammography screening in women under age 50 years," *Journal of the American Medical Association,* Vol. 259, No. 10, pp. 1512-1519.

Etchason J., Petz L., Keeler E. et al., 1995, "The cost-effectiveness of preoperative autologous blood donations," *New. England. Journal of Medicine,* Vol. 332, No. 11, pp. 719-724.

Garner T.I. and Dardis R., 1987, "Cost-effectiveness analysis of end stage renal disease treatments," *Medical Care,* Vol. 25, No. 1, pp. 25-34.

Gold M.R., Siegel J.E., Russell L.B. and Weinstein M.C. *eds.*, 1996, *Cost-effectiveness in health and medicine,* Oxford University Press, New York.

Gottlob C.A., Pellissier J.M., Wixson R.L., Stern S.H., Stulberg S.D., Chang R.W., 1995, "A cost-effectiveness analysis of total knee replacement," American Orthopedic Association Annual Resident's Conference, Pittsburgh, Pennsylvania.

Harris W.H. and Sledge C.B., 1990, "Total Hip and total knee replacement," *New England. Journal of Medicine,* Vol. 323, No. 11, pp. 725-731.

Hazen G.B., 1992, "Stochastic Trees: A new technique for temporal medical decision modeling," *Medical Decision Making,* Vol. 12, No. 3, pp. 163-178.

Hazen G.B., 1993, "Factored stochastic trees: A tool for solving complex temporal medical decision making problems," *Medical Decision Making,* Vol. 13, No. 3, pp. 227-236.

Hazen G.B., 1997, "*StoTree 1* Installation", Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, Illinois, 60208.

Hazen G.B., Hopp W.H. and Pellissier J.M., 1991, "Continuous-risk utility assessment in medical decision making," *Medical Decision Making*, Vol. 11, No. 4, pp. 294-304.

Hazen G.B., and Pellissier J.M.., 1996, "Recursive utility for stochastic trees," *Operations Research*, Vol. 44, No. 5, pp. 788-809.

Hillner B.E., Hollenberg J.P., and Pauker S.G., 1986, "Postmenopausal estrogens in prevention of osteoporosis," *American Journal of Medicine*, Vol. 80, No. 6, pp. 1115-1127.

Hollenberg J.P., 1984, "Markov cycle trees: A new representation for complex Markov processes," (abstr.). *Medical Decision Making*, Vol. 4, No. 4, p. 529.

Kavanagh B.F. and Fitzgerald Jr. R.H., 1985, "Clinical and roentgenographic assessment of total hip arthroplasty: A new hip score," *Clinical Orthopaedics and Related Research*, Vol. 193, March, pp. 133-140.

Kavanagh B.F. and Fitzgerald Jr. R.H., 1987, "Multiple revisions for failed total hip arthroplasty not associated with infection," *Journal of Bone and Joint Surgery*, Vol. 69A, No. 8, pp. 1144-1149.

Kavanagh B.F., Dewitz M.A., Ilstrup D.M., Stauffer R.N., Coventry M.B., 1989, "Charnley total hip arthroplasty with cement: fifteen-year results," *Journal of Bone and Joint Surgery*, Vol. 71A, No. 10, pp. 1496-1503.

Matchar D.B., and Pauker S.G., 1986, "Transient ischemic attacks in a man with coronary artery disease: Two strategies neck and neck," *Medical Decision Making*, Vol. 6, No. 4, pp. 239-249.

Mooney C., Mushlin A.I., and Phelps C.E., 1990, "Targeting assessments of magnetic resonance imaging in suspected multiple sclerosis," *Medical Decision Making*, Vol. 10, No. 2, pp. 77-94.

Mushlin A.I. and Fintor L., 1992, "Is screening for breast cancer cost-effective?" *Cancer*, Vol. 69, No. 7, Supplemental, pp. 1957-1962.

Orthopedic Knowledge Update, Volume IV, 1993, Rosemont, Illinois: American Academy of Orthopaedic Surgeons.

Pauker S.G., 1976, Coronary artery surgery: the use of decision analysis," *Annals of Internal Medicine*, Vol. 85, No. 1, pp. 8-18.

Pellissier J.M. and Hazen G.B., 1994, "Implementation of continuous risk utility assessment: the total hip replacement decision," *Socio-Economic Planning Sciences*, Vol. 28, No. 4, pp. 251-276.

Pellissier J.M., Hazen G.B., and Chang R.W., 1996, "A continuous risk decision analysis of total hip replacement," *Journal of the Operational Research Society*, Vol. 47, No. 6, pp. 776-793.

Plante D.A., Piccirillo J.F., and Sofferman R.A., 1987, "Decision analysis of treatment options in pyriform sinus carcinoma," *Medical Decision Making*, Vol. 7, No. 2, pp. 74-83.

Raiffa H. 1968, *Decision Analysis*, Addison-Wesley, Reading, Massachusetts.

Ransohoff D.F., Gracie W.A., Wolfenson L.B., Neuhauser D., 1983, "Prophylactic cholecystectomy or expectant management for silent gallstones," *Annals of Internal Medicine*, Vol. 99, No. 2, pp. 199-204.

Roach P.J., Fleming C., Hagen M.D., Pauker S.G., 1988, "Prostatic cancer in a patient with asymptomatic HIV infection: are some lives more equal than others?" *Medical Decision Making*, Vol. 8, No. 2, pp. 132-144.

Roberts S.D., Maxwell D.R. and Gross T.L., 1980, "Cost-effective care of end-stage real disease: A billion dollar question," *Annals of Internal Medicine*, Vol. 92, No. 2, Part 1, pp. 243-248.

Ross S.M., 1970, *Applied probability models with applications*, Holden-Day, San Francisco, California.

Ross S.M., 1989, *Introduction to probability models*, 4th ed., Academic Press, San Diego, California.

Sanzen L., Carlsson A.S., Josefsson G. and Lindberg L.T., 1988, "Revision operations on infected total hip arthroplasties," *Clinical Orthopaedics and Related Research*, Vol. 229, April, pp. 165-172.

Schulman K.A., Lynn L.A., Glick H.A., Eisenberg J.M., 1991, "Cost-effectiveness of low-dose zidovudine therapy for asymptomatic patients with human immunodeficiency virus (HIV) infection," *Annals of Internal Medicine*, Vol. 114, No. 9, pp. 798-802.

Sonnenberg A., 1985, "Comparison of different strategies for treatment of duodenal ulcer," *British Medical Journal Clinical Research Edition*, Vol. 290, April 20, pp. 1185-1187.

Steinbroker O., Traeger C. H., and Batterman R. C., 1949, "Therapeutic criteria in rheumatoid arthritis," *Journal of the American Medical Association*, Vol. 340, pp. 659-662.

Tsevat J., Eckman M.H., McNutt R.A., Pauker S.G., 1986, "Warfarin for dilated cardiomyopathy: a bloody tough pill to swallow?" *Medical Decision Making*, Vol. 9, No. 3, pp. 162-169.

van Saase J.L.C.M., 1989, "Osteoarthrosis in general population: A follow-up study of osteoarthrosis of the hip." Ph.D. dissertation, Erasmus University, Rotterdam.

Weinstein M.C., Fineberg H.V., Elstein A.S., Frazier H.S., Neuhauser D., Neutra R.R., McNeil B.J., 1980, *Clinical Decision Analysis*, W.B. Saunders Company, Philadelphia, Pennsylvania.

Weinstein M.C., and Stason W.B., 1976, *Hypertension: A Policy Perspective*, Harvard University Press, Cambridge, Massachusetts.

Weinstein M.C. and Stason W.B., 1985, "Cost-effectiveness of interventions to prevent or treat coronary heart disease," *Annals Rev. Public Health*, Vol. 6, No. 9, pp. 41-63.

Appendix: Stochastic Tree Rollback

Consider a subtree H of a stochastic tree in which an initial state x is occupied until one of several competing transitions with rates λ_i to subsequent subtrees K_i occurs (Figure 12).

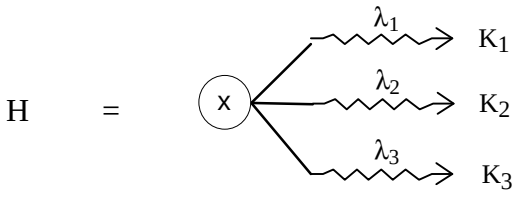

Figure 12: The stochastic tree H is a subtree of a larger tree. H begins at state x and has stochastic arcs with rates λ_i leading to further subtrees K_i.

Suppose the mean quality-adjusted duration beginning at the start of K_i is $L(K_i)$. Hazen (1992) shows that mean quality-adjusted duration L(H) beginning at x is

$$L(H) = \frac{v(x) + \sum_i l_i L(K_i)}{\sum_i l_i}. \qquad (1)$$

Mean quality-adjusted duration can be recursively evaluated in any stochastic tree by averaging the subsequent L(H) values at chance nodes and using (**1**) at stochastic nodes. If the tree contains directed cycles (repeated visits to a state) this formula will lead to iterative calculations. These equations are simple enough to be implemented on a spreadsheet.

Hazen and Pellissier [1996] define a class of utility functions which are risk-sensitive extensions of quality-adjusted duration. They may be described as follows. Define a *history* as a finite sequence $h = x_1^{t_1} x_2^{t_2} \ldots x_n^{t_n}$ of states $x_1, x_2, \ldots, x_n$ and durations t_1,

t_2, ..., t_n, where it is understood that state x_1 is first occupied for duration t_1, followed by state x_2 for duration t_2, and so on. Recursively define a utility function u(h) over histories $h = x^t k$ with initial sojourn t in state x by

$$u(h) = \int_0^t v(x)e^{-a(x)s}ds + e^{-a(x)t}u(k) \qquad (2)$$

where it is understood u(k) = 0 when k is null (i.e., when h is simply x^t). The quantity v(x) is a *quality rate* which assigns quality v(x)ds to each small interval ds spent in state x. The quantity a(x) may be thought of as a state-dependent discount rate. When h is simply a duration-t sojourn in x, (2) yields by integration

$$u(x^t) = \begin{cases} \dfrac{v(x)}{a(x)}\left(1 - e^{-a(x)t}\right) & a(x) \neq 0 \\ v(x)t & a(x) = 0 \end{cases} \qquad (3)$$

so we see that a(x) may also be interpreted as the coefficient of constant risk aversion for durations spent in state x. If a(x) is positive, then the decision maker is risk averse for uncertain lifetimes spent in state x, whereas if a(x) is negative, the decision maker is risk seeking. If a(x) is equal to zero, the decision maker is risk neutral.

The expected utility of a stochastic tree H which begins at state x and makes transitions to one of the succeeding subtrees K_1, ..., K_n (Figure 12) can be shown to satisfy the following recursive equation:

$$E[u(H)] = \frac{v(x) + \sum_i \lambda_i E[u(K_i)]}{a(x) + \sum_i \lambda_i} \qquad (4)$$

provided $a(x) + \sum_i \lambda_i > 0$ (and otherwise equal to $+\infty$). In the risk-neutral case, in which $a(x) = 0$, expected utility is equivalent to mean quality-adjusted duration with quality rates $v(x)$. Since expected utility is infinite when $a(x)$ becomes too negative, only limited degrees of risk-seeking preference may be modeled. For practically occurring degrees of risk seeking, we have not found this to be a difficulty.