DECISION ANALYSIS
Vol. 3, No. 4, December 2006, pp. 208–219
ISSN 1545-8490 | EISSN 1545-8504 | 06 | 0304 | 0208

informs

DOI 10.1287/deca.1060.0078

Parametric Sensitivity Analysis Using Large-Sample Approximate Bayesian Posterior Distributions

Gordon B. Hazen, Min Huang
Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, Illinois 60208-3119
{gbh305@lulu.it.northwestern.edu, huangmin@northwestern.edu}

When a decision analyst desires a sensitivity analysis on model parameters that are estimated from data, a natural approach is to vary each parameter within one or two standard errors of its estimate. This approach can be problematic if parameter estimates are correlated or if model structure does not permit obvious standard error estimates. Both of these difficulties can occur when the analysis of time-to-event data—known as survival analysis—plays a significant role in the decision analysis. We suggest that in this situation, a large-sample approximate multivariate normal Bayesian posterior distribution can be fruitfully used to guide either a traditional threshold proximity sensitivity analysis, or a probabilistic sensitivity analysis. The existence of such a large-sample approximation is guaranteed by the so-called Bayesian central limit theorem. We work out the details of this general proposal for a two-parameter cure-rate model, used in survival analysis. We apply our results to conduct both traditional and probabilistic sensitivity analyses for a recently published decision analysis of tamoxifen use for the prevention of breast cancer.

Key words: sensitivity analysis; probabilistic sensitivity analysis; value of information; large-sample distributions; multivariate normal approximations; Bayesian central limit theorem; cure-rate model; survival analysis; tamoxifen; breast cancer prevention
History: Received on March 13, 2006. Accepted by Robert Clemen and Robert Winkler on September 18, 2006, after 2 revisions.

1. Introduction

In a decision analysis, it is standard practice to perform sensitivity analyses on parameter estimates. When an analyst uses data to estimate an individual parameter such as a probability, the accompanying standard error can be used to guide the sensitivity analysis. The analyst may, for example, vary the parameter within one or two standard errors of its estimate. However, for more sophisticated models, an entire parameter vector may be estimated by such techniques as maximum likelihood. Due to the nature of the likelihood functions for these models, parameter estimates rarely take the simple form of an estimate with associated standard error. Moreover, parameter estimates may be correlated, in which case it does not make sense to vary them individually in a sensitivity analysis. How, then, should an analyst go about structuring a sensitivity analysis?

It is useful to discuss these difficulties in Bayesian terms. If a Bayesian posterior distribution on parameters given data is available, it would be natural to conduct a sensitivity analysis by varying parameters within high-probability regions of the posterior, or to sample from the posterior distribution in a *probabilistic sensitivity analysis* (see below). Unfortunately, model parameters may not be independent a posteriori, so it is not valid to use marginal distributions to conduct a multiway sensitivity analysis. Instead, one must sample from the multivariate posterior, or use it to determine a high-probability region. The difficulty of the latter tasks is heightened when, as is often the case, the posterior distribution has no closed form.

One approach to address these issues has been available for some time in the statistical literature—the so-called Bayesian central limit theorem (DeGroot 1968, Sen 1994), which states that for sufficiently large sample sizes, posterior distributions given data are approximately multivariate normal. Drawing samples from a multivariate normal distribution is relatively easy, and finding high-probability regions is straightforward. Using an approximate multivariate normal posterior to guide a sensitivity analysis is therefore an

attractive option, one that has not, to our knowledge, been applied in the decision analysis literature.

The purpose of this paper is to work out the details of this approach for a particular survival analysis model, the cure-rate model (Ibrahim et al. 2001). Survival analysis—the analysis of time-to-event data—has seen widespread application in engineering, economics, public health, biology, and medicine. The parameters of the cure-rate model are the probability that the failure mode of concern has been eliminated ("cured") and one or more parameters for the failure time distribution, assuming the failure mode is still present. When the failure-time distribution is exponential, the model is a simple Markov chain, and may be readily incorporated into larger Markov models. Due to data censoring, which is inevitable for survival models, the posterior distribution given failure time data has no closed form. Moreover, parameters may be highly correlated a posteriori. However, as we show, these difficulties may be bypassed because, by the Bayesian central limit theorem, the posterior distribution of these parameters is approximately multivariate normal for large sample sizes.

In this article, we address both traditional and probabilistic sensitivity analyses for the two-parameter cure-rate model with exponential failure-time distribution, or for larger models that incorporate it. We conclude this introduction below with a discussion of these types of sensitivity analysis. The remainder of the paper is organized as follows: We present specifics of the cure-rate model in §2. In §3 we describe the use of large-sample approximate normal posteriors for sensitivity analysis on cure-rate parameters. Finally, in §4 we present illustrative sensitivity analyses for a published decision analysis of tamoxifen to prevent breast cancer. We conduct a probabilistic sensitivity analysis and a one-way sensitivity analysis for a cancer survival component, and we also do an overall probabilistic sensitivity analysis for the model. Although the illustrations we present throughout arise in medical decision analysis, these methods can be used in arbitrary applications of survival models.

Sensitivity Analysis

Two types of post hoc sensitivity analysis are practiced in the decision analysis community. In a traditional threshold-proximity sensitivity analysis, once one has determined the optimal policy corresponding to one's best estimate of parameter values, one then varies parameter values across a reasonable range and observes whether any policy changes result. If policy changes occur only for parameter values far from one's best estimates, then one can feel confident in recommending the optimal policy. Otherwise, it may be necessary to improve estimates by collecting more data, or resign oneself that the optimal policy is a "close call."

In a probabilistic sensitivity analysis (Doubilet et al. 1985, Critchfield and Willard 1986), the analyst assigns probability distributions to uncertain parameters and can thereby compute or estimate as a measure of robustness the probability of a change in the optimal alternative due to variation in an arbitrary number of parameters, or alternately, the expected value of perfect information regarding any set of parameters.

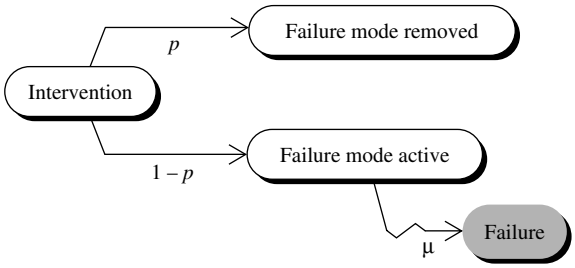
Probabilistic sensitivity analysis is popular in the medical decision analysis community. The *Guide to the Methods of Technology Appraisal* (2004) published by Britain's National Institute of Clinical Excellence states that "[p]robabilistic sensitivity analysis should be conducted on models to reflect the combined implications of uncertainty in parameters" (p. 26). The *International Society for Pharmaceutical and Outcomes Research* Task Force on Good Research Practices policy is that "[s]pecification of probability distributions for input parameters based on sampling uncertainty and/or between-study variations may be incorporated into formal probabilistic sensitivity analysis" (Weinstein et al. 2003, p. 13). The *Guidelines for Economic Evaluation of Pharmaceuticals: Canada* (1997) recommends probabilistic sensitivity analysis for situations in which "there is the possibility of important variation in the cost-effectiveness ratios or suspicion of the interdependence of variables" (p. 48).

In this article, we address both threshold proximity and probabilistic sensitivity analyses. Our approach is explicitly Bayesian, that is, we treat the cure-rate parameters as uncertain quantities, and compute approximately normal posterior distributions given survival data. We use these distributions to conduct both threshold-proximity and probabilistic sensitivity analyses.

2. The Cure-Rate Model

Figure 1 shows a generic cure-rate model formulated as a stochastic tree (Hazen 1992, 1993, 2002). This

Figure 1 **A Simple Markov Chain Depicting the Cure-Rate Model, Where p Is the Probability of Successful Intervention (Cure) and μ Is the Failure Rate, Assuming the Failure Mode Is Still Active**



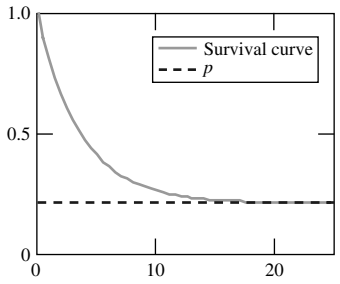
model assumes that intervention to address a failure mode of concern leads in the short term to either successful removal (or *cure*) of the failure mode with probability p, or that the failure mode remains with probability $1-p$. Whether the failure mode remains active is unobserved unless failure occurs later, and the failure rate if the mode is active is μ. The advantage of such a simple model is its easy formulation as a Markov chain, and the simplicity of incorporating it into larger decision analysis models.

In the cure-rate model, the survival function $S(t)$, equal to the probability of no failure for at least duration t, is given by

$$S(t) = p + (1-p)e^{-\mu t}$$

and is graphed in Figure 2. It slopes downward from 1 and has asymptote equal to the probability p of cure.

Figure 2 **The Survival Function $S(t) = p + (1-p)e^{-\mu t}$ of the Cure-Rate Model**



Parameter Estimation

Survival data typically involves *censoring*, in which the failure time of a unit remains unobserved due to mechanisms independent of the failure process. Let t_i be the time of censoring of unit i in our data set, and suppose the data set has n_D observed failures, among which the average failure time is $\bar{t}_D$. Under the cure-rate model, the likelihood function for parameters (p, μ) for the n observations is given by

$$L(p,\mu) \propto (1-p)^{n_D}\mu^{n_D}e^{-\mu n_D \bar{t}_D} \prod_{\substack{i:\text{subject } i \\ \text{censored}}} (p + (1-p)e^{-\mu t_i}),$$

where $\propto$ denotes proportionality up to a constant that involves the censoring mechanism. The *maximum-likelihood estimate* $(\hat{p}, \hat{\mu})$ is the value of (p, μ) that maximizes $L(p, \mu)$.

In this model, if the true values of p and μ satisfy $0 < p < 1$ and $\mu > 0$, then the maximum-likelihood estimates (MLEs) are consistent, i.e., they converge to the real parameters as sample size approaches infinity (DeGroot 1968, Sen 1994). Due to the awkward form of the likelihood function, there is no simple formula for MLEs, which must be computed by nonlinear optimization methods.

Example 1. Table 1 shows breast cancer survival data for women 50–54 years of age from the National Cancer Institute's Surveillance, Epidemiology and End Results (or SEER) database for the years 1990 to 2001.

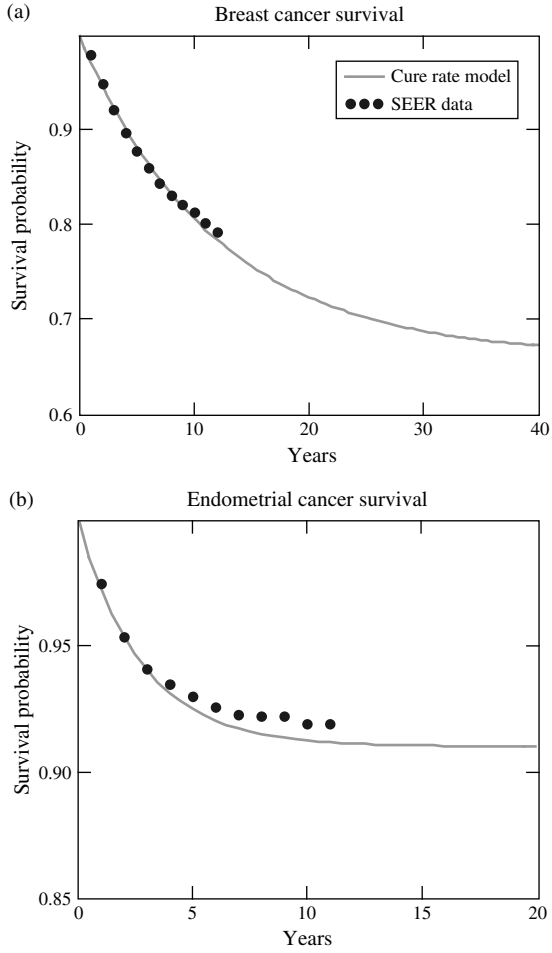
Here, $n_D = 2{,}101$ and $\bar{t}_D = 3.05$. The MLE under the cure rate model is $(\hat{p}, \hat{\mu}) = (0.66, 0.084)$. The estimated survival function $S(t) = \hat{p} + (1-\hat{p})e^{-\hat{\mu}t}$ is graphed in Figure 3(a) along with the original SEER data.

Example 2. Figure 3(b) shows SEER endometrial cancer survival data from 1990 to 2001 for women

Table 1 **Breast Cancer Survival Data for Women 50–54 Years of Age from the SEER Database**

Years	Alive at start	Died	Lost to follow-up
<1	19,089	357	2,058
1 ≤2	16, 674	490	2,051
2-<3	14,133	396	1,936
3-<4	11,801	279	1,765
4-<5	9,757	193	1,711
5-<6	7,853	139	1,497
6-<7	6,217	103	1,283
7-<8	4,831	67	1,213
8-<9	3,551	34	1,002
9-<10	2,515	24	933
10-<11	1,558	14	815
11-<12	729	5	724

Figure 3 **(a) Theoretical Breast Cancer Survival Function for the Cure-Rate Model with MLE $(\hat{p}, \hat{\mu}) = (0.66, 0.084)$, Compared to the SEER Data and (b) Theoretical Endometrial Cancer Survival Function for the Cure-Rate Model with MLE $(\hat{p}, \hat{\mu}) = (0.91, 0.36)$, Compared to the SEER Data**



50–54 years of age, and the corresponding cure-rate survival curve, again using maximum-likelihood estimates.

As we can see, despite its simple form, the cure rate model using the maximum-likelihood estimates for p and μ does give a reasonable fit to cancer survival data.

3. Sensitivity Analysis

Although the MLE $(\hat{p}, \hat{\mu})$ is a good estimate of model parameters, it is merely an estimate. If one wishes to conduct a sensitivity analysis on p and μ, how should one proceed? In what follows, we draw on the Bayesian central limit theorem discussed in the introduction. The basic idea is to obtain an approximate posterior probability distribution for (p, μ) and use it to guide a sensitivity analysis.

Bayesian Posterior Distribution of Parameters (p, μ) We treat the parameters p, μ as random variables with specified prior distributions. If we assume uniform noninformative prior, then the posterior distribution of (p, μ) is proportional to the likelihood function. Therefore, the posterior density function of parameters (p, μ) given observed data is

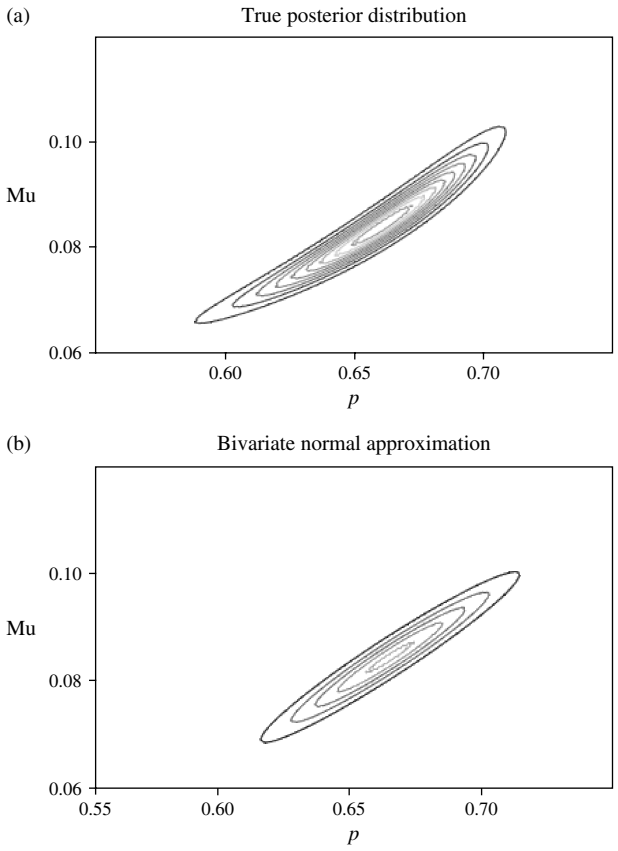
$$f(p, \mu \mid \text{Obs}) \propto (1-p)^{n_D} \mu^{n_D} e^{-\mu n_D \bar{t}_D} \cdot \prod_{\substack{i:\ \text{subject } i \\ \text{censored}}} (p + (1-p)e^{-\mu t_i}). \quad (1)$$

However, its awkward analytical form makes this posterior distribution difficult to use for sensitivity analysis. Due to the censoring terms in this distribution, p and μ are correlated, and as we will see shortly, the correlation can be very high. Therefore, it is not advisable to vary each one independently in a one-way sensitivity analysis. However, to conduct a two-way sensitivity analysis, we need to form, say, a 95% credible region for (p, μ). The awkward form of the posterior distribution makes this inconvenient as well. Failing this, one can manually select a discrete set of points (p, μ) from a contour plot of the posterior distribution (Figure 4 below), an approach that may be adequate for some purposes.

There are analogous difficulties for probabilistic sensitivity analysis, where one would wish to generate random (p, μ) pairs from the posterior distribution. However, the awkward form of the posterior makes standard techniques difficult to apply. One possibility would be to use Markov chain Monte Carlo (MCMC) techniques. However, MCMC is inherently less robust than analytic statistical methods, and there are still many remaining issues about the behavior of existing methods in real applications (Gilks et al. 1996). Moreover, we prefer sensitivity analysis techniques that are easier to implement. The approach we present can be carried out on a spreadsheet.

The solution we propose—the Bayesian central limit theorem applied to the cure-rate model—takes

Figure 4 **Contour Graphs of the Joint Posterior Distribution of Probability of Cure of Breast Cancer (p) and Mortality Rate If Not Cured (μ) Given the SEER Data**



Notes. Diagram (a) shows the true posterior distribution and (b) graphs the large-sample approximate bivariate normal distribution.

the following form. See DeGroot (1968) and Sen (1994) for the general theory needed to derive this result. We provide a proof in the appendix available in the online supplement to this article (http://da.pubs.informs.org/online-supp.html).

Theorem 1. *If the parameters (p, μ) in the cure-rate model have strictly positive prior density over the region $0 < p < 1$, $\mu > 0$, then the posterior distribution of (p, μ) is for large samples approximately a multivariate normal distribution with mean vector $(\hat{p}, \hat{\mu})$ and covariance matrix $(-H)^{-1}$, where $(\hat{p}, \hat{\mu})$ is the maximum-likelihood estimate (MLE), and H is the Hessian of the log-posterior evaluated at $(\hat{p}, \hat{\mu})$.*

More specifically, the Hessian H is the matrix of second partial derivatives with respect to p and μ of the log of the posterior (1):

$$H = \begin{pmatrix} -\dfrac{n_D}{(1-p)^2} - \displaystyle\sum_i \frac{(1-e^{-\mu t_i})^2}{S(t_i)^2} & \displaystyle\sum_i \frac{t_i e^{-\mu t_i}}{S(t_i)^2} \\ \displaystyle\sum_i \frac{t_i e^{-\mu t_i}}{S(t_i)^2} & -\dfrac{n_D}{\mu^2} + p(1-p)\displaystyle\sum_i \frac{t_i^2 e^{-\mu t_i}}{S(t_i)^2} \end{pmatrix} \tag{2}$$

$$S(t_i) = p + (1-p)e^{-\mu t_i}.$$

The required calculations for the MLE and covariance matrix can be easily done on a spreadsheet having an optimization add-in.

Example 3 (Continuation of Example 1). Given the breast cancer survival data for women 50–54 years of age from Table 1, the MLE for (p, μ) is obtained in Example 1: $\hat{p} = 0.66$, $\hat{\mu} = 0.084$. From (2), the negative of the Hessian matrix evaluated at $(\hat{p}, \hat{\mu})$ is

$$-H = \begin{pmatrix} 2.14 \times 10^4 & -6.34 \times 10^4 \\ -6.34 \times 10^4 & 2.01 \times 10^5 \end{pmatrix}$$

and the approximate covariance matrix is

$$-H^{-1} = \begin{pmatrix} 7.15 \times 10^{-4} & 2.26 \times 10^{-4} \\ 2.26 \times 10^{-4} & 7.63 \times 10^{-5} \end{pmatrix}.$$

We conclude that (p, μ) has an approximate bivariate normal posterior distribution with mean vector $(0.66, 0.084)^T$, and covariance matrix $-H^{-1}$. In other words, the approximate standard deviations of p and μ are

$$\sigma_p = 0.0267 \qquad \sigma_\mu = 0.0087, \tag{3}$$

and the approximate correlation between p and μ is

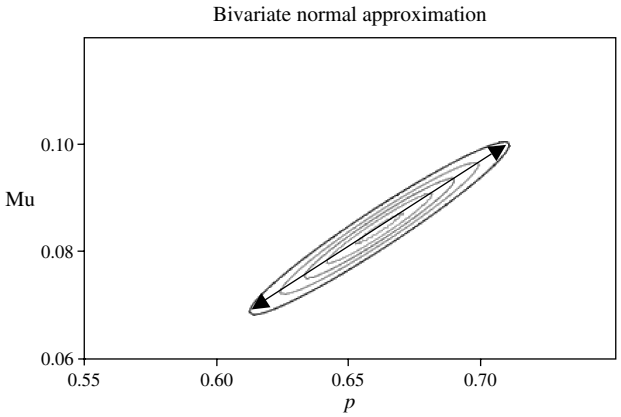
$$\rho_{p\mu} = 0.967. \tag{4}$$

We may compare contour graphs for this approximate bivariate normal distribution for (p, μ) and the true posterior distribution (1), and the result is shown in Figure 4. As we can see, the approximation is reasonably close.

Transformation of Variables

For moderate sample sizes, the approximating bivariate normal distribution over (p, μ) may place unacceptably large probability mass outside the feasible region $[0, 1] \times [0, \infty)$, a situation that may be revealed from plots like Figure 4b, or by examining the approximate marginal distributions of p and μ. The standard recourse is then to reparameterize to unbounded

Figure 5 Contour Graph of the Approximate Bivariate Normal Posterior Distribution p and μ from the Breast Cancer Cure-Rate Model Given the SEER Data



Note. The double arrow denotes the principal component.

parameters via the logit and log transformations, that is,

$$\theta_1 = \ln\left(\frac{p}{1-p}\right) \qquad \theta_2 = \ln \mu.$$

The approximating posterior over $\theta = (\theta_1, \theta_2)$ is then bivariate normal with mean equal to the maximum-likelihood estimate $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2) = (\ln(\hat{p}/(1-\hat{p})), \ln \hat{\mu})$ and covariance matrix equal to the inverse of the negative of a new Hessian matrix $H_{new}(\hat{\theta})$ evaluated at $\hat{\theta}$. The latter is given by

$$H_{new}(\hat{\theta}) = DH(\hat{p}, \hat{\mu})D,$$

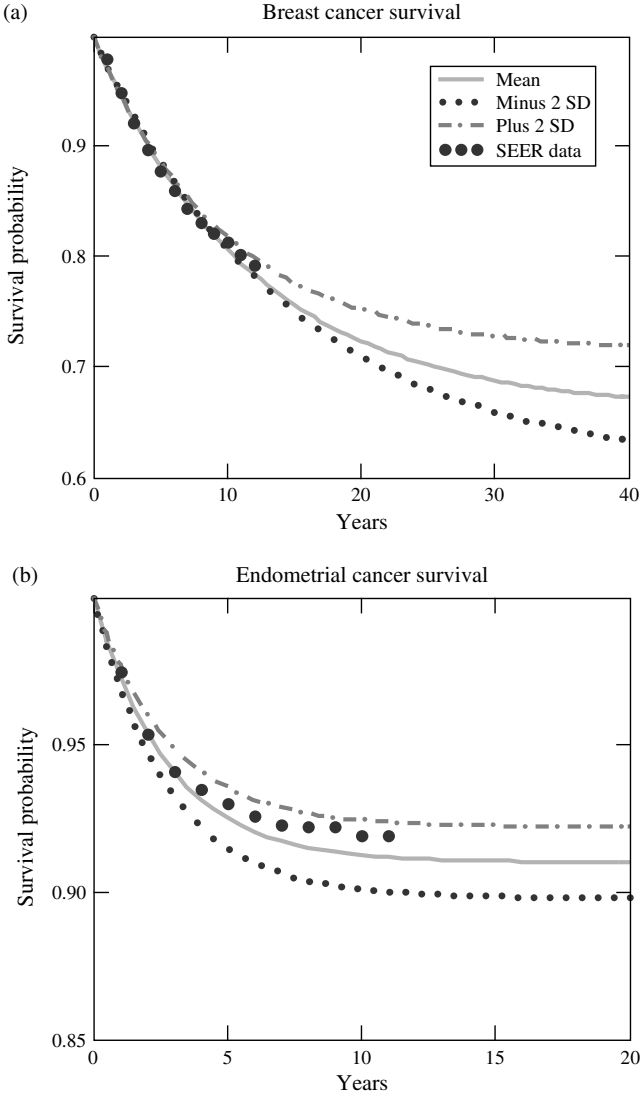
where D is the diagonal matrix with entries $\hat{p}(1-\hat{p})$ and $\hat{\mu}$ in that order, and $H(\hat{p}, \hat{\mu})$ is the Hessian (2) evaluated at the maximum-likelihood estimate $(\hat{p}, \hat{\mu})$.

One-Way Sensitivity Analysis

We can use principal components of the approximate normal posterior density to do a one-way sensitivity analysis. We vary p and μ simultaneously two standard deviations along the principal component, as indicated in Figure 5 by the double arrow. We discuss the details for finding principal components in the appendix available in the online supplement (http://da.pubs.informs.org/online-supp.html). The notion of *synthetic variables* introduced for one-way sensitivity analysis by Reilly (2000) is similar in spirit to this approach, although the specifics differ.

Example 4 (Continuation of Example 3). The resulting variation in breast cancer survival is given in Figure 6a. Survival duration is sensitive to variations of (p, μ) about their MLE only for durations in excess of SEER observations. Variation in endometrial cancer survival is shown in Figure 6b.

Figure 6 (a) Breast Cancer Survival Probabilities post Diagnosis Under the Approximate Bivariate Normal Posterior of (p, μ) (b) Endometrial Cancer Survival Probabilities post Diagnosis, Again Under the Approximate Bivariate Normal Posterior of (p, μ)



Note. The upper and lower dashed lines represent two standard deviations along the largest principal component in (p, μ).

Decision Analysis 3(4), pp. 208–219, © 2006 INFORMS

Probabilistic Sensitivity Analysis

The approximate bivariate normal posterior can also be used to conduct a probabilistic sensitivity analysis. To generate random (p, μ) values from the bivariate normal distribution, we use the result that if (p, μ) is bivariate normal with means m_p and m_μ, variances σ_p^2 and σ_μ^2, and covariance $\sigma_{p\mu}$, then (e.g., Tong 1990, Anderson 2003)

- the marginal distribution of p is normal with mean m_p and variance σ_p^2
- the conditional distribution of μ given p is normal with mean $m_{\mu|p} = m_\mu + ((p - m_p)/\sigma_p)\rho_{p\mu}\sigma_\mu$ and variance $\sigma_{\mu|p}^2 = \sigma_\mu^2(1 - \rho_{\mu p}^2)$.

Using the values (3) and (4) from Example 3, we have

$$m_p = 0.66 \qquad \sigma_p = 0.0267.$$

$$m_{\mu|p} = 0.084 + 0.316 \cdot (p - 0.66)$$

$$\sigma_{\mu|p} = 0.00223.$$

Therefore, we can generate random variates p, μ in two steps:

1. Generate p from a normal distribution with mean 0.66 and standard deviation 0.0267.

2. Given this value of p, generate μ from a normal distribution with mean $0.084 + 0.316(p - 0.66)$ and standard deviation 0.00223.

4. An Illustrative Example

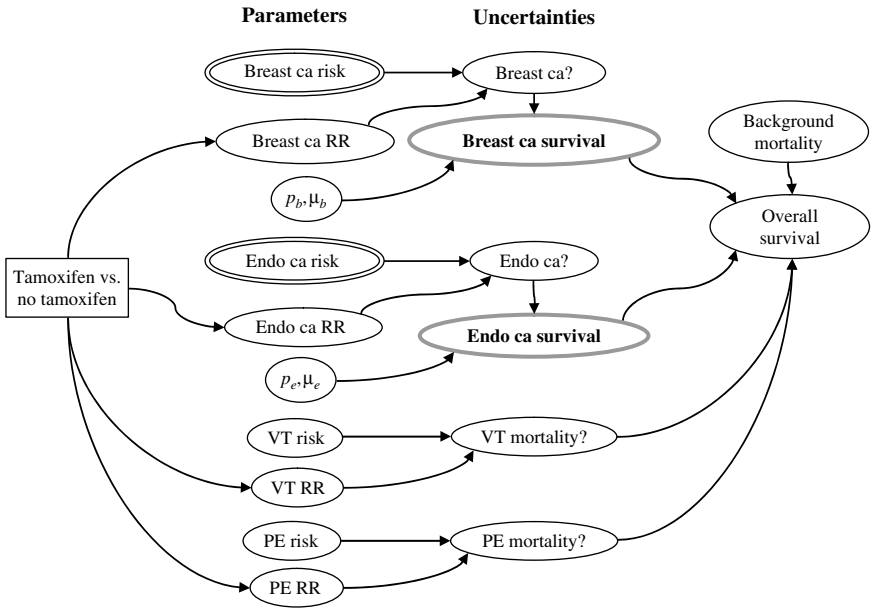
Models for postdiagnosis cancer survival are components of many medical decision and cost-effectiveness analyses. For instance, they are important components of any cancer-screening model, and may be incorporated into cancer prevention models. One example of the latter is an analysis by Col et al. (2002) of the decision to use tamoxifen for breast cancer prevention. At the time, tamoxifen was regarded as an effective hormonal therapy against established breast cancer, and it was therefore natural to investigate its potential as a preventive agent in healthy subjects. Because tamoxifen is associated with potentially life-threatening side effects such as endometrial cancer, venous thrombosis, and pulmonary embolism, there are both risks and benefits of its use. The goal of Col and colleagues was to augment the earlier analysis of Gail et al. (1999) to account for differences in mortality and timing of occurrences, as well as alternate time-dependent effects of tamoxifen on breast and endometrial cancer risk.

As a specific application of the sensitivity analysis techniques introduced above, we reproduced the Col et al. (2002) model and conducted sensitivity analyses based on large-sample normal approximations to cure rate models. However, since the time of our reanalysis, results of the Study of Tamoxifen and Raloxifene (STAR) trials (e.g., Vastag 2006) have appeared, which show that the drug raloxifene produces the same breast cancer risk reduction as tamoxifen in high-risk postmenopausal women, with less risk of endometrial cancer and other side effects. Raloxifene therapy is now regarded as the preferred treatment for breast cancer risk reduction in high-risk women. The Col et al. (2002) decision analysis is therefore no longer relevant for breast cancer prevention. We present it here for purely illustrative purposes.

Col et al. (2002) developed a Markov model to estimate the impact of tamoxifen use for the prevention of breast cancer in healthy women. Figure 7 shows an influence diagram we constructed to depict this problem. It has not only breast cancer survival but also endometrial cancer survival as components. We modified the Col et al. model by using cure-rate models for breast cancer survival (parameters p_b, μ_b) and endometrial cancer survival (parameters p_e, μ_e). These models, with cancer incidence added, are shown in Figure 8.

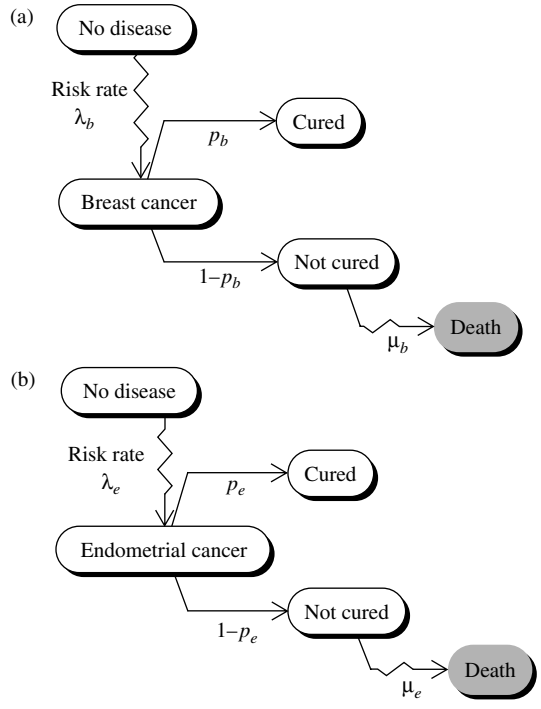
In Table 2, we list all probability, rate, and relative risk parameters used in the model. Col et al. (2002) performed analyses for four different levels of breast cancer risk, combined with four different levels of endometrial cancer risk, including zero risk for women with no uterus due to hysterectomy. In this example, we do an analysis for 50-year-old women in only one of the Col et al. categories, namely, those who are at fourfold higher and twofold higher risk for developing breast cancer and endometrial cancer, respectively. Thus, following Col et al. (2002), we set the risk rate λ_{b0} of breast cancer without tamoxifen to be 0.0086/year, corresponding to a five-year risk of approximately 4.19%, four times higher than average. At the same time, we set endometrial cancer risk

Figure 7 An Influence Diagram for the Decision to Use Tamoxifen for Breast Cancer Prevention, Based on Col et al. (2002)



Note. RR = relative risk. Tamoxifen reduces breast cancer risk (RR < 1), but increases risks of endometrial cancer (endo ca), venous thrombosis (VT), and pulmonary embolism (PE) (RR > 1). We modified the Col et al. model by using cure-rate models for breast cancer (breast ca) survival (parameters p_b, μ_b) and endometrial cancer survival (parameters p_e, μ_e).

Figure 8 Cancer Incidence and Cure-Rate Models for Breast Cancer Survival post Diagnosis (a) and Endometrial Cancer Survival post Diagnosis (b)



rate λ_{e0} to be 0.0015/year, which doubles the average risk of approximately 0.00076/year (Fisher et al. 1998). Col et al. considered two scenarios associated with five years of tamoxifen use: a breast cancer benefit lasting 10 years and endometrial cancer risk lasting five years; and a breast benefit of 15 years along with endometrial risk lasting 10 years. We reproduced only the first of these two scenarios.

Baseline Results

In our baseline analysis, we used the specified breast cancer risk and endometrial cancer risk and the means of posterior distributions for the other 12 parameters. This analysis and the subsequent sensitivity analysis were conducted in Microsoft Excel, using the *Stotree* software written by one of the authors (Hazen 2002, see also http://users.iems.northwestern.edu/~hazen/). We found only a small life-expectancy gain of 0.41 month for tamoxifen. For the identical scenario, Col et al. (2002) found a 3.8 month gain. Col et al. found a range of zero to five months life-expectancy gain depending on breast and endometrial cancer risk levels, so we do not regard the difference between 0.41 and 3.8 months as large.

Table 2 Risk Rate, Relative Risk (RR), Probability and Mortality Rate Parameters in the Col et al. Model

Variable	Description	Baseline value
Breast ca risk λ_{b0}	Risk rate of breast cancer without tamoxifen	0.0086/year
Breast ca RR	RR of breast cancer with tamoxifen vs. without	0.4936
Endo ca risk λ_{e0}	Risk rate of endometrial cancer without tamoxifen	0.0015/year
Endo ca RR	RR of endometrial cancer with tamoxifen vs. without	4.0132
VT risk	Risk rate of venous thrombosis without tamoxifen	0.0009/year
VT RR	RR of venous thrombosis with tamoxifen vs. without	1.7159
PE risk	Risk rate of pulmonary embolism without tamoxifen	0.0003/year
PE RR	RR of pulmonary embolism with tamoxifen vs. without	3.2258
p_b	Probability of cure for breast cancer	0.6610
μ_b	Mortality rate of breast cancer if not cured	0.0844/year
p_e	Probability of cure for endometrial cancer	0.9102
μ_e	Mortality rate of endometrial cancer if not cured	0.3576/year
VT mortality	Short-term mortality of venous thrombosis	0.0091
PE mortality	Short-term mortality of pulmonary embolism	0.3380

Notes. The risk rate of breast cancer is set to be fourfold higher than average risk, the risk rate of endometrial cancer is set to double average risk, and other baseline values listed are posterior means of parameters.

One-Way Sensitivity Analysis for p_b, μ_b

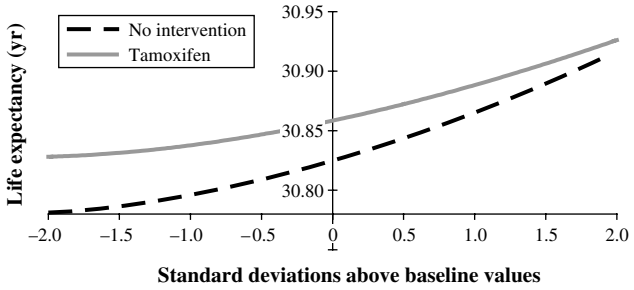
We conducted a one-way sensitivity analysis for parameters p_b, μ_b of the breast cancer survival model. Setting all other parameters equal to their baseline values, we vary p_b and μ_b, simultaneously, two standard deviations along the principal component as discussed in Example 3.

The principal component, as shown in Figure 5, is $(p_b, \mu_b) = (0.66, 0.084) + \alpha \cdot (2.27 \times 10^{-2}, 8.49 \times 10^{-3})$, where α is the number of standard deviations along the principal component. We found life expectancy with and without tamoxifen for standard deviations $-2 \le \alpha \le 2$, shown in Figure 9. The optimal decision of using tamoxifen does not change along the principal component.

Probabilistic Sensitivity Analysis for p_b, μ_b

We investigated the sensitivity of the optimal decision of using tamoxifen to the uncertainty of parameters in the breast cancer survival component. That is, a probabilistic sensitivity analysis was performed jointly on

Figure 9 Life Expectancy with and Without Tamoxifen Use, When (p_b, μ_b) Varies Within Two Standard Deviations from Its Baseline Value Along the Principal Component of Its Approximate Posterior Distribution



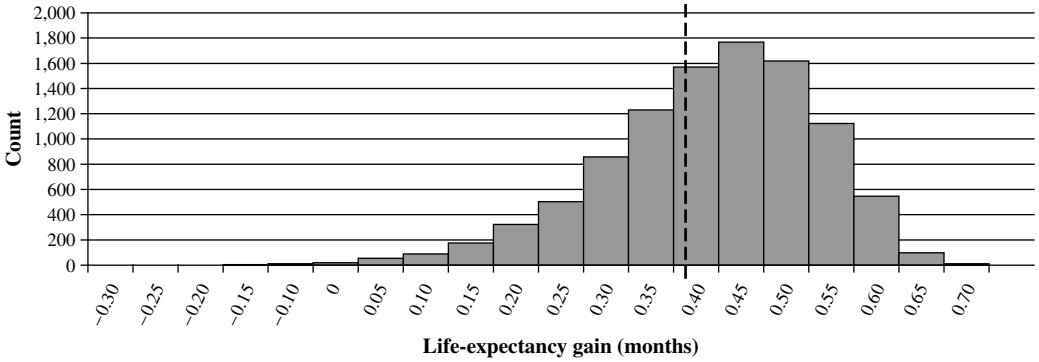
p_b, μ_b, using the method discussed in Example 3, and setting all other parameters equal to their baseline values.

Based on 10,000 Monte Carlo iterations (see Figure 10), tamoxifen has a 99.66% (s.e. 0.058%) probability of optimality, and the expected value of perfect information on p_b, μ_b equal to 0.0058 (s.e. 0.0013) days of life. Thus, the decision of tamoxifen is almost always optimal, and the expected value of perfect information on the parameters is extremely small. The optimality of tamoxifen is therefore insensitive to these two parameters. This concurs with the conclusion of one-way sensitivity analysis, that the optimality of tamoxifen is very insensitive to p_b and μ_b.

Overall Probabilistic Sensitivity Analysis

We also conducted a probabilistic sensitivity analysis jointly on the 12 probability, rate, and efficacy parameters (including the four parameters in the two cancer survival components) discussed in Table 2, leaving the cancer risks λ_{b0} and λ_{e0} specified as before. First we assigned distributions to all the uncertain parameters. We have already discussed the distributions for the four parameters in cancer survival components. For other model parameters, we use the method of constructing large-sample approximate Bayesian posterior distributions summarized in Hazen and Huang (2006), obtaining approximately transformed normal posteriors based on data provided by Fisher et al. (1998), Dalen and Alpert (1975), and Hillner et al. (1992). Our Monte Carlo simulation takes random draws for all 12 parameters from their posterior distributions and estimates the probability of decision change and the expected value of perfect information.

Figure 10 The Histogram of Life-Expectancy Gain by Using Tamoxifen, Using 10,000 Iterations of a Probabilistic Sensitivity Analysis on the Parameters p_b and μ_b



Note. Baseline life-expectancy gain is shown as a dashed vertical line.

Based on 10,000 Monte Carlo iterations (see Figure 11), tamoxifen has a 55.0% (s.e. 0.50%) probability of optimality, and the expected value of perfect information on all 12 uncertain probability, rate, and efficacy parameters equal to 11.0 (s.e. 0.18) days of life.

Sensitivity Analysis Conclusions

It is easy to lose track of the purpose of these calculations. Tamoxifen therapy is preferred—it has a small expected life benefit of 0.41 months = 12.3 days. Sensitivity analysis addresses the question of how sensitive this conclusion is to parameter uncertainty. Sensitivity to uncertainty in the breast cancer cure-rate parameters p_b, μ_b is very small. Sensitivity to uncertainty in all parameters combined is more ambiguous: There is a substantial chance (45%) that a tamoxifen recommendation could be wrong due to errors in parameter estimates, leading us to suspect that we should gather more information on these parameters before making such a recommendation. The information-value calculation belies this suspicion, however—the expected value of knowing all true parameter values is only 11 days of life. This quantity is equal to the chance (45%) that tamoxifen is suboptimal times an expected gain in life (24.4 days), given that tamoxifen is declined when it is suboptimal. Therefore, even if further research showed that tamoxifen was not beneficial, the life-expectancy gain (24.4 days) by reversing this recommendation would be small. (This contradiction between information value and optimality probability as measures of sensitivity arises frequently in applications—see Felli and Hazen 1998, 1999 for more details.) Overall, recommending tamoxifen is risky (45% chance of being

Figure 11 The Histogram of Life-Expectancy Gain by Using Tamoxifen, Using 10,000 Iterations of a Probabilistic Sensitivity Analysis on All Rate and Probability Parameters

Note. Baseline life-expectancy gain is shown by the vertical dashed line.

wrong), but there is little downside to making that recommendation, as at most an expected 11 days of life would be gained by doing further research that could reverse this recommendation.

Conclusion

We have presented large-sample Bayesian methods for conducting traditional and probabilistic sensitivity analyses on the parameters of the cure-rate model for survival analysis. The cure-rate model can be formulated easily as a Markov chain, and is therefore simple to incorporate into larger decision analysis models. Our Bayesian sensitivity analysis methods are simple enough to be performed entirely on a spreadsheet.

The cure-rate model is not strictly necessary for the tamoxifen analysis we present—it might be possible to directly use the data points in Figure 3 as estimates of survival probabilities versus time, which could be used to directly compute the mean survival advantage for tamoxifen. There are several problems with this approach: (a) Survival probabilities would still need to be extrapolated to time periods beyond the 12 years of survival data, and for this purpose, some assumptions, if not the cure-rate assumptions, would still have to be made; (b) the associated Markov model would be nonstationary, a surmountable difficulty, but still an inconvenience; and (c) the associated sensitivity analysis on the survival probability estimates would be much more complicated. Clearly, the survival probabilities, one for each time step, are not independent estimates, so it is not acceptable to vary each of them independently in a sensitivity analysis. The cure-rate model provides a built-in correlation between these estimates, and makes sensitivity analysis (relatively) easy.

The Bayesian methods we recommend are based on large-sample normal approximations. For univariate approximations, sample sizes of 30 to 50 are usually adequate, but bivariate accuracy may be less, and we are aware of no firm guidelines. Fortunately, it is easy to construct contour plots of the otherwise awkward true posterior distribution for cure-rate parameters (see our Figure 4), and this may be compared with contours of the approximate normal posterior if there are doubts about sample size.

The cure-rate model we have used has a constant mortality rate for the uncured. More sophisticated cure-rate models may allow a time-dependent mortality rate. This results in a nonstationary Markov model that is less convenient for our purposes. Nevertheless, large-sample posterior distributions for parameters in these models may be developed in much the same way we have done here. In general, for any model in which parameter estimates are not accompanied by standard errors, the use of large-sample normal approximations for posterior distributions over these parameters appears to be a viable procedure for conducting sensitivity analyses.

An online supplement to this paper is available at http://da.pubs.informs.org/online-supp.html.

References

Anderson, T. W. 2003. *An Introduction to Multivariate Statistical Analysis*, 3rd ed. Wiley, New York.

Berenson, M. L., D. M. Levine, M. Goldstein. 2001. *Intermediate Statistical Methods and Applications*. Prentice-Hall, Inc., Englewood Cliffs, NJ.

Canadian Coordinating Office for Health Technology Assessment. 1997. *Guidelines for Economic Evaluation of Pharmaceuticals: Canada*, 2nd ed. Canadian Coordinating Office for Health Technology Assessment (CCOHTA), Ottawa, Canada.

Col, N. F., R. J. Goldberg, R. K. Orr, J. K. Erban, J. M. Fortin, R. T. Chlebowski. 2002. Survival impact of tamoxifen use for breast cancer risk reduction: Projections from a patient-specific Markov model. *Medical Decision Making* **22**(5) 386–393.

Critchfield, G. G., K. E. Willard. 1986. Probabilistic analysis of decision trees using Monte Carlo simulation. *Medical Decision Making* **6**(2) 85–92.

Dalen, J. E., J. S. Alpert. 1975. Natural history of pulmonary embolism. *Prog. Cardiovasc. Dis.* **17**(4) 259–270.

DeGroot, M. H. 1968. *Optimal Statistical Decisions*. McGraw Hill, New York.

Doubilet, P., C. B. Begg, M. C. Weinstein, P. Braun, B. J. McNeil. 1985. Probabilistic sensitivity analysis using Monte Carlo simulation. *Medical Decision Making* **5**(2) 157–177.

Felli, J. C., G. B. Hazen. 1998. Sensitivity analysis and the expected value of perfect information. *Medical Decision Making* **18**(1) 95–109.

Felli, J. C., G. B. Hazen. 1999. Do sensitivity analyses really capture problem sensitivity? An empirical analysis based on information value. *Risk, Decision Policy* **4** 79–98.

Fisher, B., J. P. Costantino, D. L. Wickerham, C. K. Redmond, M. Kavanah, W. M. Cronin, V. Vogel, A. Robidoux, N. Dimitrov, J. Atkins, M. Daly, S. Wieand, E. Tan-Chiu, L. Ford. 1998. Tamoxifen for prevention of breast cancer: Report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *J. National Cancer Inst.* **90**(18) 1371–1388.

Gail, M. H., J. P. Costantino, J. Bryant, R. Croyle, L. Freedman, K. Helzlsouer, V. Vogel. 1999. Weighing the risks and benefits of tamoxifen treatment for preventing breast cancer. *J. National Cancer Inst.* **91**(21) 1829–1846.

Gilks, W. R., S. Richardson, D. J. Spiegelhalter. 1996. *Markov Chain Monte Carlo in Practice*. Chapman & Hall, London, UK.

Hazen, G. B. 1992. Stochastic trees: A new technique for temporal medical decision modeling. *Medical Decision Making* **12**(3) 163–178.

Hazen, G. B. 1993. Factored stochastic trees: A tool for solving complex temporal medical decision models. *Medical Decision Making* **13**(3) 227–236.

Hazen, G. B. 2002. Stochastic trees and the StoTree modeling environment: Models and software for medical decision analysis. *J. Medical Systems* **26**(5) 399–413.

Hazen, G., M. Huang. 2006. Large-sample Bayesian posterior distributions for probabilistic sensitivity analysis. *Medical Decision Making*. **26**(9) 512–534.

Hillner, B., J. T. Philbrick, D. N. Becker. 1992. Optimal management of suspected lower-extremity deep vein thrombosis: An evaluation with cost assessment of 24 management strategies. *Arch. Intern. Med.* **152**(1) 165–175.

Ibrahim, J. G., M.-H. Chen, D. Sinha. 2001. *Bayesian Survival Analysis*. Springer, New York.

National Institute for Clinical Excellence. 2004. *Guide to the Methods of Technology Appraisal*. Abba Litho Sales Limited, London, UK.

Reilly, T. 2000. Sensitivity analysis for dependent variables. *Decision Sci.* **31**(3) 551–572.

Sen, P. K., J. M. Singer. 1994. *Large Sample Methods in Statistics: An Introduction with Applications*. CRC Press, New York.

Tong, Y. L. 1990. *The Multivariate Normal Distribution*. Springer-Verlag, New York.

Vastag, B. 2006. Raloxifene prevails in STAR trial, may face easier road to acceptance than previous drugs. *J. National Cancer Inst.* **98**(11) 733–735.

Weinstein, M. C., B. O'Brien, J. Hornberger, J. Jackson, M. Johannesson, C. McCabe, B. R. Luce. 2003. Principles of good practice for decision analytic modeling in health-care evaluation: Report of the ISPOR task force on good research practices-modeling studies. *Value Health* **6**(1) 9–17.