

# Medical Decision Making

<http://mdm.sagepub.com/>

---

## **Cohort Decomposition for Markov Cost-Effectiveness Models**

Gordon Hazen and Zhe Li

*Med Decis Making* 2011 31: 19 originally published online 8 June 2010

DOI: 10.1177/0272989X10368999

The online version of this article can be found at:

<http://mdm.sagepub.com/content/31/1/19>

---

Published by:



<http://www.sagepublications.com>

On behalf of:



[Society for Medical Decision Making](http://www.smdm.org)

**Additional services and information for *Medical Decision Making* can be found at:**

**Email Alerts:** <http://mdm.sagepub.com/cgi/alerts>

**Subscriptions:** <http://mdm.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

>> [Version of Record](#) - Feb 4, 2011

[OnlineFirst Version of Record](#) - Jun 8, 2010

[What is This?](#)

# Cohort Decomposition for Markov Cost-Effectiveness Models

Gordon Hazen, DPhil, Zhe Li, MS

*Cohort analysis is a widespread tool for computing expected costs and quality-adjusted life years (QALYs) in Markov models for medical cost-effectiveness analyses. Although not always explicitly identified, such models commonly have multiple simple factors, or components. In these, a health state consists of a multiple component vector, one component for each factor, and arbitrary combinations of components are possible. The authors show here that when the model does not assume any probabilistic dependence among these factors, then a standard cohort analysis may be decomposed into several independent cohort analyses, one for each factor, and the results may be combined to produce desired expected costs and QALYs. These single-factor cohort*

*analyses are not only simpler but also computationally more efficient. The authors derive the appropriate formulas for this cohort decomposition in discrete time and give several examples of their use based on published cost-effectiveness analyses. Explicitly identifying the simple factors of which a model is composed allows these factors to be portrayed graphically. Graphical depiction of the simple factors that comprise a model reduces model complexity, makes model formulation easier and more transparent, and thereby facilitates peer inspection and critique. **Key words:** Markov models; mathematical models and decision analysis, cost-effectiveness analysis; women's health. (*Med Decis Making* 2011;31:19–34)*

Markov modeling has become an accepted tool for medical decision and cost-effectiveness analyses, and cohort analysis has become the standard procedure for solving such models to obtain expected costs, life years, and quality-adjusted life years.<sup>1–4</sup> In a cohort analysis, a homogeneous cohort of individuals is tracked through succeeding states in accordance with state transition probabilities specified by the Markov model. Cohort analysis is usually simple enough to be done on a spreadsheet.

The purpose of this article is to point out and provide a partial solution to what we believe is an underreported shortcoming of Markov modeling—namely, the proliferation of states that can arise when the model has multiple *factors* or *components*. In such models, a health state consists of

a multiple component vector, one component for each factor, and arbitrary combinations of components are possible. Examples are widespread. For instance, Col and colleagues<sup>5</sup> use a Markov model to investigate the survival impact of tamoxifen in breast cancer risk reduction. In so doing, they implicitly consider 5 factors—namely, incidence, progression, and mortality related to breast cancer; incidence, progression, and mortality related to endometrial cancer; incidence and mortality of venous thrombosis; incidence and mortality of pulmonary embolism; and background mortality due to other causes. Anderson and colleagues<sup>6</sup> recently investigated preventive policies for *BRCA* + women, including factors involving breast cancer, ovarian cancer, endometrial cancer, pulmonary embolism, cataracts, and background mortality. Three decades ago, Tsevat and colleagues<sup>7</sup> analyzed the use of warfarin for dilated cardiomyopathy, incorporating factors of systemic embolism, pulmonary embolism, systemic hemorrhage, and background mortality. (See also Hazen<sup>8</sup> for a description of the factors in this model.) More examples could be cited. Although these investigators do not explicitly portray their models as multifactorial, all have this structure.

Received 3 November 2008 from the Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, Illinois. Revision accepted for publication 25 January 2010.

Address correspondence to Gordon Hazen, DPhil, Northwestern University, McCormick School of Engineering and Applied Science, Evanston, IL 60208-3119; e-mail: hazen@iems.northwestern.edu.

DOI: 10.1177/0272989X10368999

The difficulty is that multifactor models, whether explicitly recognized or not, lead quickly to state proliferation. For instance, consider a model with 3 factors, each one simple and involving only 3 nonfatal states. If combinations of states in different factors are all allowable, then there are  $3 \cdot 3 \cdot 3 + 1 = 28$  combinations of states (27 nonfatal plus the dead state). This is a problem for two reasons: first, if cohort analysis is performed in a spreadsheet, then the modeler has to deal with 28 columns. Second, keeping track of the number of possible transitions that can occur with 3 factors plus background mortality operating simultaneously can be confusing for a modeler and a prohibitively daunting black box for a peer reviewer.

Modelers commonly cope with state proliferation by identifying the most important possible state combinations and restricting state transitions to these states. For example, Schousboe and colleagues<sup>9</sup> perform a cost-effectiveness analysis of alendronate therapy on osteopenic postmenopausal women. They consider 5 different types of fractures, which in principle should allow  $2^5 = 32$  state combinations corresponding to 5 factors each at 2 possible levels. However, in addition the no-fracture state, Schousboe and colleagues consider not 32 but only 5 states corresponding to a single fracture type, plus one other state corresponding to the combination of the worst 2 possible fracture types. As a consequence, the corresponding Markov model “knows” only the most recent fracture type and cannot account for, say, increased risk of a second hip fracture when there has been an intervening distal forearm fracture. These authors used a patient-level simulation model rather than a Markov model, and we suspect that this issue, as well as the related problem of state proliferation, contributed to that modeling choice.

The approach we present here as an alternate solution to state proliferation draws on the observation that in many cases, factors in a multifactor model are *probabilistically independent*—transition probabilities in one factor do not depend on the states of the other factors. This need not always hold, but we show that when it does, one can perform a *cohort decomposition*—that is, one may decompose a large Markov cohort analysis into separate cohort analyses, one for each factor, and combine the results to obtain expected costs, lifetimes, and quality-adjusted life years (QALYs) as needed. For instance, if as above there are 3 probabilistically independent factors with 3 nonfatal states each, then 3 smaller cohort analyses suffice, each with

$1 + 3 = 4$  states, for a total of 12 columns in a spreadsheet as opposed to the 28 columns required for the original cohort analysis just mentioned. Moreover, each of the 4-state analyses need consider only transitions that occur within its own factor and not others, simplifying the model formulation and making it more open to peer critique.

Cohort decomposition is not a panacea—factors might not be probabilistically independent. There is, however, one factor that can almost always be taken to be independent of any other factors—namely, *background mortality*—and cohort decomposition is therefore usually possible with respect to this factor.

This article is structured as follows. In the next section, we give general formulas and several simple examples of cohort decomposition, illustrating how expected costs, lifetimes, and QALYs may be computed. Then, we provide computational results for a realistically sized Markov cohort decomposition based on Schousboe and colleagues’ cost-effectiveness analysis<sup>9</sup> of alendronate therapy mentioned above. The last section concludes with a discussion. Derivations of results are provided in an appendix.

## DISCRETE-TIME COHORT DECOMPOSITION

In this section, we introduce cohort decomposition and give several simple examples of its use in discrete-time Markov models to compute expected costs and QALYs. We classify costs and QALYs as either *ongoing* (i.e., produced by the continuous occupancy of a health state) or *transition induced* (i.e., produced as *tolls* by a transition from one health state to another). We describe cohort decomposition calculations for each type of cost and QALY.

### Ongoing QALYs

We illustrate the computation of ongoing QALYs with the following example, adopted from Roach and colleagues.<sup>10</sup> These authors analyze the choice between radiation therapy and radical prostatectomy for stage B prostatic cancer in a 58-year-old man with asymptomatic HIV infection. The Markov cycle tree used in their analysis is shown in Figure 1. The states in the figure represent the long-term potential consequences of radical prostatectomy: the cancer state represents prostate cancer recurrence, the AIDS state represents the development of

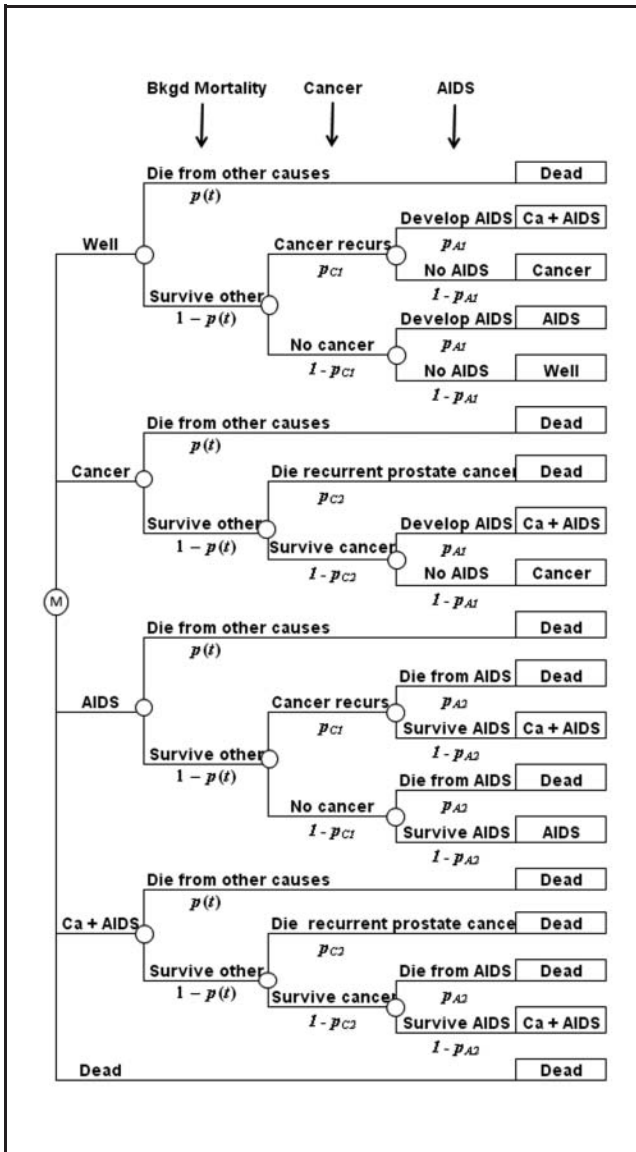


Figure 1 Markov cycle tree for the analysis by Roach and colleagues<sup>10</sup> of the treatment of recurrent prostate cancer.

full-blown AIDS, and the Ca + AIDS state represents their combination.

Notice in Figure 1 that the probabilities of developing or dying from AIDS are the same regardless of cancer status and vice versa. Moreover, other-cause survival is independent of cancer and AIDS status. This makes possible a much more compact multifactor depiction of the same model, as is shown in Figure 2.

To implement a cohort decomposition, our first task is to perform a separate cohort analysis for

each factor. The cohort analysis for the cancer factor would calculate the expected numbers  $n_{\text{No Cancer}}(t)$ ,  $n_{\text{Cancer}}(t)$ ,  $n_{\text{Dead}}(t)$  of individuals in the no cancer, cancer, and dead states at month  $t$ , respectively, using the initial values

$$n_{\text{No Cancer}}(0) = N_{\text{Ca}}, n_{\text{Cancer}}(0) = 0, n_{\text{Dead}}(0) = 0$$

and the one-step transition formula

$$n_y(t) = \sum_x n_x(t - \Delta t) p_{xy}, \quad (1)$$

where  $x, y \in \{\text{No Cancer}, \text{Cancer}, \text{Dead}\}$ ,  $\Delta t$  is the cycle length (1 year in this case), and  $p_{xy}$  is the per cycle transition probability from state  $x$  to state  $y$ . The initial cohort size  $N_{\text{Ca}}$  here is arbitrarily chosen to give an easier interpretation to the numerical results (Figure 3) and does not affect final per person values.

The same formula would be used to obtain the expected numbers  $n_{\text{No AIDS}}(t)$ ,  $n_{\text{AIDS}}(t)$ ,  $n_{\text{Dead}}(t)$  of individuals in the no AIDS, AIDS, and dead states at month  $t$  in the AIDS factor. The resulting values for the cancer factor and the AIDS factor are shown in Figure 3. Cohort analysis for the background mortality factor is a special case of (1), namely, with  $n_{\text{Alive}}(0) = N_{\text{B}}$ ,

$$n_{\text{Alive}}(t) = n_{\text{Alive}}(t - \Delta t)(1 - p(t)).$$

Only the resulting survival probabilities  $S_{\text{B}}(t) = n_{\text{Alive}}(t)/N_{\text{B}}$  are shown in Figure 3.

As can be seen in Figure 1 but also calculated as follows from Figure 2, there are  $1 + 2 \cdot 2 \cdot 1 = 5$  state combinations in this model—namely, well (no cancer or AIDS), cancer, AIDS, Ca + AIDS, and dead. Roach and colleagues<sup>10</sup> assign quality coefficients 1.0, 0.6, 0.5, 0.3, and 0 to these states. Observe that the quality coefficient 0.3 assigned to Ca + AIDS is the product of the coefficients  $v_{\text{C}} = 0.6$  for cancer and  $v_{\text{A}} = 0.5$  for AIDS. Because of this, the corresponding quality coefficients  $v(y_{\text{C}}, y_{\text{A}}, y_{\text{B}})$  in the factored model of Figure 2 obey the product rule

$$v(y_{\text{C}}, y_{\text{A}}, y_{\text{B}}) = v_{\text{C}}(y_{\text{C}}) \cdot v_{\text{A}}(y_{\text{A}}) \cdot v_{\text{B}}(y_{\text{B}}). \quad (2)$$

where  $y_{\text{C}} \in \{\text{No Cancer}, \text{Cancer}, \text{Dead}\}$  is a state in the cancer factor,  $y_{\text{A}} \in \{\text{No AIDS}, \text{AIDS}, \text{Dead}\}$  is a state in the AIDS factor,  $y_{\text{B}} \in \{\text{Alive}, \text{Dead}\}$  is a state in the background mortality factor, and  $v_{\text{C}}, v_{\text{A}}, v_{\text{B}}$  are the corresponding quality coefficients in each factor.

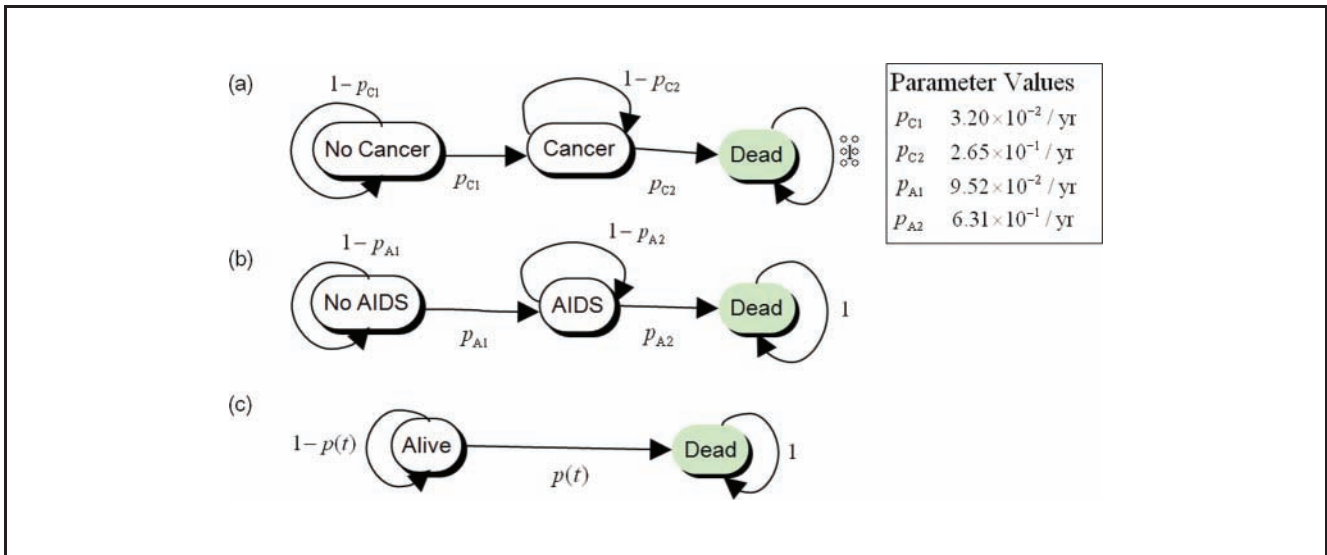


Figure 2 Factored model of cancer recurrence and AIDS incidence in the presence of background mortality. (a) State transition diagram for cancer incidence and death. (b) State transition diagram for AIDS incidence and death. (c) State transition diagram for background mortality.

	Cancer				AIDS				Bkgd Mortality		
Health quality	1	0.6	0		1	0.5	0		Mortality Rate /10,000	Survival Prob.	QALY Rate/ person
Year	No Cancer	Cancer	Death	E[QALY Rate]/ person	No AIDS	AIDS	Death	E[QALY Rate]/ person			
0	10000	0	0	1.0000	10000	0	0	1.0000		1.0000	0.5
1	9680	320	0	0.9872	9048	952	0	0.9524	99	0.9901	0.9038
2	9371	545	85	0.9697	8187	1212	601	0.8793	107	0.9796	0.7873
3	9071	700	229	0.9491	7408	1226	1366	0.8021	116	0.9682	0.6745
4	8781	804	415	0.9264	6703	1157	2140	0.7282	126	0.9561	0.5730
5	8500	872	628	0.9023	6065	1064	2870	0.6598	136	0.9431	0.4843
6	8228	912	859	0.8776	5488	970	3542	0.5973	148	0.9293	0.4080
7	7965	934	1101	0.8525	4966	880	4154	0.5406	160	0.9146	0.3427
8	7711	941	1349	0.8275	4493	797	4710	0.4892	173	0.8988	0.2872
9	7464	938	1598	0.8027	4066	721	5213	0.4426	188	0.8821	0.2402
10	7225	928	1847	0.7782	3679	653	5668	0.4005	203	0.8644	0.2005
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
38	2908	399	6693	0.3148	224	40	9737	0.0244	1894	0.0952	0.0002
39	2815	386	6799	0.3047	202	36	9762	0.0220	2051	0.0775	0.0002
40	2725	374	6901	0.2950	183	33	9784	0.0199	2221	0.0621	0.0001
									Total		6.3331

Figure 3 Cohort decomposition for the calculation of expected overall quality-adjusted life year (QALY) in the model of Figure 2. The 3 boxes show the cohort analysis for the cancer, AIDS, and background mortality factors, respectively. The arrows indicate the product calculation from equation (3).

Equation (2) and the probabilistic independence of the 3 factors guarantee that expected ongoing QALYs can be computed using cohort decomposition. This is illustrated in Figure 3 and can be described as follows. The expected quality accrual rate (QR) per person  $E[QR_{Ca}(t)]$  at time  $t$  in the cancer (Ca) factor is

$$E[QR_{Ca}(t)] = \frac{1}{N_{Ca}} (1 \cdot n_{No\ Cancer}(t) + v_C \cdot n_{Cancer}(t)),$$

where  $n_{No\ Cancer}(t)$ ,  $n_{Cancer}(t)$  are the numbers of people in the no cancer and cancer states at time  $t$ , respectively. This formula reflects the fact that  $n_{No\ Cancer}(t)$  individuals accrue QALYs at rate 1 per year and  $n_{Cancer}(t)$  individuals accrue QALYs at rate  $v_C$  per year.

Similarly, if  $N_A$  is the initial AIDS factor cohort size, the expected quality accrual rate per person  $E[QR_A(t)]$  at time  $t$  in the AIDS (A) factor is

$$E[QR_A(t)] = \frac{1}{N_A} (1 \cdot n_{No\ AIDS}(t) + v_A n_{AIDS}(t)).$$

( $N_{Ca}$  and  $N_A$  are both equal to 10,000 in Figure 3, although there is no need for them to be the same.) The expected quality accrual rate per person  $E[QR_B(t)]$  at time  $t$  in the background mortality (B) factor is just the survival probability, that is,

$$E[QR_B(t)] = \frac{1}{N_B} n_{Alive}(t) = S_B(t),$$

where  $N_B = 10,000$  is the background mortality cohort size. Because of the multiplicative relation (2) and probabilistic independence, overall expected quality accrual rate  $E[QR(t)]$  at time  $t$  is then the product of these 3 quality accrual rates:

$$E[QR(t)] = E[QR_{Ca}(t)] \cdot E[QR_A(t)] \cdot S_B(t). \quad (3)$$

as indicated in Figure 3, with product values appearing in the final column. Half-cycle corrections<sup>3</sup> have been used in each factor. Overall ongoing QALYs per person are obtained as the discounted sum

$$E[\text{Ongoing QALY per person}] = \sum_t (1+r)^{-t} E[QR(t)] \cdot \Delta t$$

equal to approximately 6.33 years in Figure 3.

Both cohort decomposition and conventional cohort analysis are discrete-time approximations to the true underlying continuous-time model that would obtain by letting the step size  $\Delta t$  approach

zero. As approximations, they do not necessarily yield the same QALY results. However, the two methods approach equality as the cycle length  $\Delta t$  approaches zero, and in continuous time, they would be exactly equal.

For a small model such as this one, the spreadsheet column requirements for the factored calculations are actually slightly greater than for a unified cohort analysis—advantages for the factored approach only accrue for larger models (see our discussion below). Nevertheless, the greater simplicity in cohort calculations for each factor is evident and is a direct consequence of the greater simplicity in model formulation (Figure 2 v. Figure 1).

In general, let  $v(x)$  be the quality coefficient for a state  $x = (x_1, \dots, x_n)$ . Then cohort decomposition can be applied when  $v(x)$  is expressible as the product

$$v(x) = \prod_i v_i(x_i) \quad (4)$$

of quality coefficients  $v_i(x_i)$  from each factor  $i$ . In addition to the Roach and colleagues<sup>10</sup> model just discussed, assumption (4) has seen use in several modeling efforts.<sup>11–14</sup> When it holds, expected ongoing QALYs per person may be obtained via cohort decomposition as

$$E[\text{Ongoing QALY per person}] = \sum_t (1+r)^{-t} E[QR(t)] \cdot \Delta t$$

with

$$E[QR(t)] = \prod_i E[QR_i(t)]$$

and

$$E[QR_i(t)] = \frac{1}{N_i} \sum_{x_i} v_i(x_i) n_{x_i}(t),$$

where the cohort numbers  $n_{x_i}(t)$  are obtained from a cohort analysis of factor  $i$  only,  $N_i$  is the initial cohort size in factor  $i$ ,  $E[QR_i(t)]$  is the expected quality accrual rate per person in factor  $i$  at time  $t$ , and  $E[QR(t)]$  is the overall expected quality accrual rate per person at time  $t$ .

### Ongoing Costs

To compute ongoing costs, a similar method is possible, which requires, in contrast to (4), only that costs add across factors. We consider the following

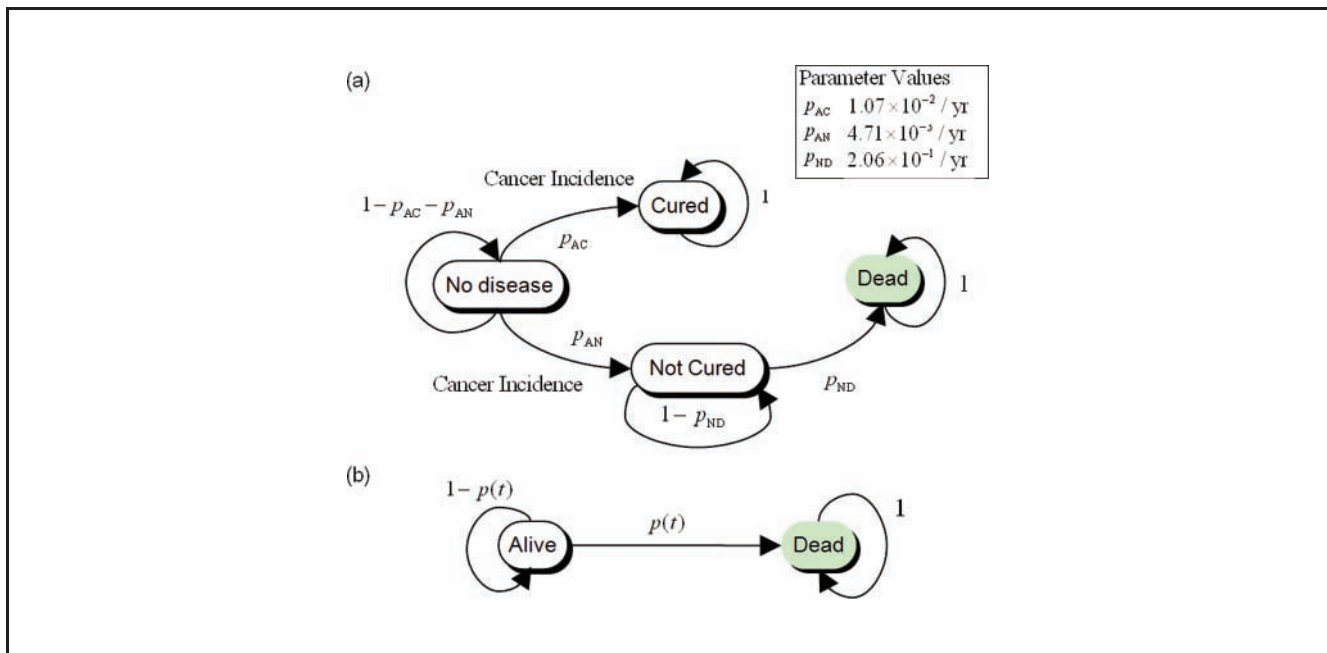


Figure 4 Factored model of cancer incidence and treatment in the presence of background mortality. (a) State transition diagram for cancer risk and treatment in discrete time. (b) State transition diagram for background mortality with mortality probability  $p(t)$  in period  $t$ .

simple example involving 2 factors: cancer incidence/treatment and background mortality, as depicted in Figure 4. In the cancer factor (Figure 4a), an individual begins in the no disease state. If cancer incidence occurs, treatment is initiated, and the individual is either cured or not. The combination of cancer incidence and either not cured or cured occurs with yearly probabilities  $p_{AC}$  and  $p_{AN}$ , respectively. Individuals in the cured state remain there, but individuals in the not cured state are subject to a yearly cancer mortality probability of  $p_{ND}$ . The probabilities shown are representative for women at high risk of ovarian cancer and were obtained from Anderson and colleagues<sup>6</sup> and the Surveillance, Epidemiology, and End Results (SEER) database.<sup>15</sup>

Figure 4b depicts the second factor in this model, background mortality. Here an individual begins in the alive state and has a yearly probability  $p(t)$  of death during year  $t$ . These probabilities would be obtained from standard life tables, and values for a 30-year-old woman are partially shown in the spreadsheet in Figure 5.

There are 2 types of ongoing costs in this model. Women at high risk for ovarian cancer would undergo surveillance in the no disease state, at cost rate  $c_S = \$3601/\text{year}$ . Women in the states of cured and not cured bear an ongoing treatment cost

$c_T = \$10,055/\text{year}$  (the same cost because whether a woman is cured is not observable). We obtained these estimates from Anderson and colleagues' model<sup>6</sup> of ovarian cancer. We postpone the consideration of one-time costs associated with cancer incidence and mortality. Figure 5 depicts a cohort analysis for this ovarian cancer factor. Here we assume that at  $t=0$ , a cohort of size  $N=10,000$  begins in the no disease state.

Once the cohort numbers  $n_x(t)$  are calculated, one can compute what we call *expected cost accrual rates*, in this case, the expected surveillance cost rate (SCR) per person  $E[\text{SCR}(t)]$  and the expected treatment cost rate (TCR) per person  $E[\text{TCR}(t)]$  at each year  $t$ :

$$E[\text{SCR}(t)] = \frac{1}{N} c_S n_{\text{No Disease}}(t). \quad (5)$$

$$E[\text{TCR}(t)] = \frac{1}{N} c_T (n_{\text{Cured}}(t) + n_{\text{Not Cured}}(t)). \quad (6)$$

These expected cost rates are shown in columns 2 to 4 of the spreadsheet in Figure 5.

Background survival probabilities are calculated as in the previous section. Once we have performed a cohort analysis for each factor, we can combine

them in the following way to obtain discounted expected costs per person:

$$\begin{aligned} & E[\text{Surveillance cost per person}] \\ &= \sum_t (1+r)^{-t} S_B(t) \cdot E[\text{SCR}(t)] \cdot \Delta t, \end{aligned} \quad (7)$$

$$\begin{aligned} & E[\text{Treatment cost per person}] \\ &= \sum_t (1+r)^{-t} S_B(t) \cdot E[\text{TCR}(t)] \cdot \Delta t. \end{aligned} \quad (8)$$

Here  $r$  is the discount rate, and  $(1+r)^{-t}$  is the discount factor for year  $t$  ( $r=3\%/y$  in Figure 5). Intuitively, to get, say, the total expected surveillance cost, we have to sum  $\Delta t$  times the discounted expected surveillance cost rates, but the latter must be multiplied by the survival probability  $S_B(t)$  because we have surveillance costs only for individuals still alive.

The calculations (7) and (8) are displayed in the final 2 columns of the spreadsheet in Figure 5. Expected surveillance cost is \$70,553, and expected treatment cost is \$48,875 for each patient during her lifetime. As in the case of QALY calculation, these are only approximately equal to the expected costs one would obtain by performing a unified cohort analysis—again, the 2 methods approach equality as the cycle length  $\Delta t$  approaches zero.

In general, for multifactor computation, the expected cost accrual rate (CR) per person in factor  $i$  at time  $t$  is given by

$$E[\text{CR}_i(t)] = \frac{1}{N_i} \sum_{x_i} c_{x_i} n_{x_i}(t),$$

where  $c_x$  is the cost rate for each unit of time spent in state  $x$ , and  $N_i$  is the cohort size for factor  $i$ . To obtain the overall expected ongoing cost, one multiplies by survival probabilities  $S_j(t)$  in other factors and forms the discounted sum

$$E[\text{Ongoing cost per person for factor } i] = \sum_t (1+r)^{-t} E[\text{CR}_i(t)] \cdot \Delta t \cdot \prod_{j \neq i} S_j(t).$$

The survival function  $S_j(t)$  at time  $t$  for factor  $j \neq i$ , equal to the probability that individuals in factor  $j$  are not dead at time  $t$ , is given by

$$S_j(t) = \frac{1}{N_j} \sum_{j \neq \text{Death}} n_j(t).$$

We presented the cohort decomposition of Figure 5 for pedagogical purposes, as it does not reduce the number of columns in the spreadsheet compared to a standard cohort analysis. Nevertheless, cohort decomposition is still simpler in the sense of allowing the modeler to consider cancer transitions and background mortality transitions separately. For realistic multifactor models, there can, however, be dramatic savings in column requirements.

### Cost Tolls

We return to the example of Figure 4. In addition to the ongoing costs discussed there, there are cost tolls associated with state transitions. These are one-time costs incurred over a short period of time and associated with a change in health state. For instance, Anderson and colleagues<sup>6</sup> estimate treatment costs  $c_I = \$55,823$  associated with ovarian cancer incidence, treatment costs  $c_M = \$38,743$  associated with ovarian cancer mortality, and costs  $c_B = \$28,787$  associated with other-cause mortality.

The expected value of these cost tolls can be calculated via cohort decomposition in the following way. In the cancer factor, the expected cancer incidence (CaI) cost toll rate (CTR) per person at month  $t$  is given by

$$E[\text{CTR}_{\text{CaI}}(t)] = c_I \cdot \frac{1}{N_{\text{Ca}}} n_{\text{No Disease}}(t) \cdot (p_{AC} + p_{AN}),$$

where  $N_{\text{Ca}}$  is the cohort size in the cancer factor. Here the term  $n_{\text{No Disease}}(t) \cdot (p_{AC} + p_{AN})$  is the expected number in the no disease state at time  $t$  who experience cancer incidence in the next time interval  $\Delta t$ . Similarly, the expected cancer mortality (CaM) cost toll rate per person at month  $t$  is given by

$$E[\text{CTR}_{\text{CaM}}(t)] = c_M \frac{1}{N_{\text{Ca}}} \cdot n_{\text{Not Cured}}(t) \cdot p_{ND}.$$

These 2 calculations are illustrated in the spreadsheet of Figure 6. These costs are incurred only when the background mortality factor is in the alive state. To obtain expected overall cost, we form a discounted sum, including a survival term for background mortality. Thus, expected discounted cost toll per person for cancer incidence is given by

$$E[\text{CT}_{\text{CaI}}] = \sum_t (1+r)^{-t} E[\text{CTR}_{\text{CaI}}(t)] \cdot S_B(t). \quad (9)$$

Expected discounted cost toll per person for cancer mortality is given by



	Ovarian Cancer						Bkgd Mortality			
Cost/ year	\$3,601	\$10,055	\$10,055	\$0						
Year	No Disease	Cured	Not Cured	Death	E[SCR] /person	E[TCR] /person	Mortality rate /10,000	Survival Prob	Surveillance Cost/person	Treatment Cost/person
0	10000	0	0	0	\$3,601	\$0	5	1.0000	\$1,801	\$0
1	9846	107	47	0	\$3,545	\$155	6	0.9995	\$3,440	\$151
2	9694	213	84	10	\$3,491	\$298	6	0.9989	\$3,287	\$281
3	9544	317	112	27	\$3,437	\$431	7	0.9983	\$3,140	\$394
4	9397	419	134	50	\$3,384	\$556	7	0.9976	\$2,999	\$493
5	9252	520	151	78	\$3,332	\$674	8	0.9968	\$2,865	\$580
6	9109	619	163	109	\$3,280	\$786	9	0.9960	\$2,736	\$656
7	8969	717	172	142	\$3,230	\$894	10	0.9951	\$2,613	\$723
8	8830	813	179	178	\$3,180	\$997	11	0.9942	\$2,496	\$783
9	8694	908	184	215	\$3,131	\$1,097	12	0.9931	\$2,383	\$835
10	8560	1001	187	253	\$3,082	\$1,194	13	0.9920	\$2,275	\$881
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
68	3473	4535	86	1906	\$1,251	\$4,646	1942	0.1183	\$20	\$74
69	3420	4572	84	1924	\$1,231	\$4,682	2119	0.0974	\$16	\$59
70	3367	4609	83	1941	\$1,212	\$4,718	2311	0.0788	\$6	\$23
								Total	\$70,553	\$48,875

Figure 5 Spreadsheet cohort decomposition used to calculate expected ongoing costs. Separate cohort analyses for the ovarian cancer and background mortality factors are outlined on the left and middle. As suggested by the arrows, the values on the right are obtained by multiplying the corresponding cells along with the discount factor, according to equations (7) and (8). Summing the results gives the overall discounted costs, \$70,553 and \$48,875.

$$E[CT_{CaM}] = \sum_t (1+r)^{-t} E[CTR_{CaM}(t)] \cdot S_B(t). \quad (10)$$

Calculation of these quantities is shown in Figure 6. The resulting expected cost tolls are \$16,728 and \$3085, respectively.

The calculation of expected cost toll for background mortality proceeds similarly. The expected background mortality cost toll rate per person is given by

$$E[CTR_B(t)] = c_B \frac{1}{N_B} n_{\text{Alive}}(t) \cdot p(t) = c_B S_B(t) p(t),$$

where we have used the fact that  $n_{\text{Alive}}(t)/N_B$  is the background survival probability  $S_B(t)$ . This cost is incurred only when the cancer factor is not in dead state. To obtain expected overall background mortality toll, we again form a discounted sum, including a survival term for cancer mortality. Thus, expected discounted cost toll per person for background mortality is given by

$$E[CT_B] = \sum_t (1+r)^{-t} E[CTR_B(t)] \cdot S_{Ca}(t).$$

Here the survival probability  $S_{Ca}(t)$  for the cancer factor is given by

$$S_{Ca}(t) = \frac{1}{N_{Ca}} (n_{\text{No Disease}}(t) + n_{\text{Cured}}(t) + n_{\text{Not Cured}}(t)).$$

These calculations (not shown in Figure 6) give  $E[CT_B] = \$5363$ .

In general, for expected cost toll calculation, we have

$$\begin{aligned} E[\text{Cost toll/person for factor } i] \\ = \sum_t (1+r)^{-t} E[CTR_i(t)] \cdot \prod_{j \neq i} S_j(t) \end{aligned}$$

with

$$E[CTR_i(t)] = \frac{1}{N_i} \sum_x n_x(t) \sum_y c_{xy} p_{xy}(t).$$

Here  $c_{xy}$  is the cost of transition  $x \rightarrow y$  in factor  $i$ ,  $p_{xy}(t)$  is the probability of this transition in  $[t, t + \Delta t)$ ,  $N_i$  is the cohort size for factor  $i$ , and  $S_j(t)$  is the factor  $j$  survival probability at time  $t$ .

Year	Ovarian Cancer				E[CTR <sub>CaI</sub> ] /person		Bkgd Mortality		Cost Toll for cancer incidence	Cost Toll for cancer mortality
	No Disease	Cured	Not Cured	Death	E[CTR <sub>CaI</sub> ] /person	E[CTR <sub>CaM</sub> ] /person	Mortality rate /10,000	Survival Prob		
0	10000	0	0	0			5	1.0000		
1	9846	107	47	0	\$854	\$0	6	0.9995	\$414	\$0
2	9694	213	84	10	\$841	\$38	6	0.9989	\$791	\$35
3	9544	317	112	27	\$828	\$67	7	0.9983	\$756	\$61
4	9397	419	134	50	\$815	\$90	7	0.9976	\$722	\$79
5	9252	520	151	78	\$802	\$107	8	0.9968	\$690	\$92
6	9109	619	163	109	\$790	\$120	9	0.9960	\$659	\$100
7	8969	717	172	142	\$778	\$130	10	0.9951	\$629	\$105
8	8830	813	179	178	\$766	\$138	11	0.9942	\$601	\$108
9	8694	908	184	215	\$754	\$143	12	0.9931	\$574	\$109
10	8560	1001	187	253	\$742	\$147	13	0.9920	\$548	\$108
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
68	3473	4535	86	1906	\$301	\$70	1942	0.1183	\$5	\$1
69	3420	4572	84	1924	\$297	\$68	2119	0.0974	\$4	\$1
70	3367	4609	83	1941	\$292	\$67	2311	0.0788	\$1	\$0
								Total	\$16,141	\$2,992

Figure 6 A spreadsheet to calculate the expected cost tolls for cancer incidence and mortality using cohort decomposition. The first 2 boxes show the cohort analysis for the factors of ovarian cancer and background mortality, respectively. As indicated by the arrows, the values in the third box are the product of the corresponding cells in the first 2 boxes. Overall cost tolls \$16,141 and \$2992 are the discounted sums formed according to equations (9) and (10).

**Morbidity Tolls**

Similar to cost tolls, there could be morbidity or QALY tolls associated with transitions, such as when an individual suffers short-term disutility due to disease incidence.

The computation of QALY tolls is also very similar to that of cost tolls, except that we must distinguish between absolute and proportional tolls. Suppose the transition  $x \rightarrow y$  in factor  $i$  carries a QALY toll  $\delta_{xy}$ . Then the overall absolute toll is given by

$$E[\text{QALY toll/person for factor } i] = \sum_t (1+r)^{-t} E[\text{QTR}_i(t)] \cdot \prod_{j \neq i} S_j(t),$$

where

$$E[\text{QTR}_i(t)] = \frac{1}{N_i} \sum_x n_x(t) \sum_y \delta_{xy} p_{xy}(t).$$

In some situations, it may be more natural to assume that the overall toll produced by a transition  $x_i \rightarrow y_i$

in factor  $i$  is proportional to the product  $\prod_{j \neq i} v_j(x_j)$

QALY coefficients in other factors in states  $x_j$ . This might occur when the toll  $\delta_{x_i y_i}$  in factor  $i$  is meant to approximate a decrement  $\Delta v_i$  of quality in factor  $i$  over a short period of time  $\Delta t$ , that is,  $\delta_{x_i y_i} = \Delta v_i \Delta t$ , and it is judged that quality coefficients should multiply across factors. In this case,

$$E[\text{QALY toll/person for factor } i] = \sum_t (1+r)^{-t} E[\text{QTR}_i(t)] \cdot \prod_{j \neq i} E[\text{QR}_j(t)].$$

In some models, it may happen that QALY coefficients  $v_j(x_j)$  are equal to 1 in all other factors  $j$ . In this case, absolute and proportional tolls are identical.

**Computational Advantages**

In terms of computational effort, if we consider a model with  $f$  independent factors (not including background mortality) each with  $s$  nondeath states, then naive cohort analysis requires  $1 + s^f$  columns

in a spreadsheet, whereas cohort decomposition requires  $1 + (1 + s)f$  columns, and as we mentioned earlier, this begins to be a big advantage for cohort decomposition when the number  $f$  of factors grows to 3 or more, even if the number  $s$  of states remains at 2 or 3 per factor, as is typical. Regardless of the software (spreadsheet or other), computational work in cohort analysis is proportional to the number of state transitions. Supposing the number of transitions in a factor with  $s$  nondeath states is roughly  $s$  also, this results in  $s^f$  transitions in the overall model on which naive cohort analysis is applied, compared to  $s \cdot f$  transitions in cohort decomposition, again a big advantage for cohort decomposition when the number of factors grows to 3 or more.

## COHORT DECOMPOSITION FOR FRACTURE RISK ASSESSMENT

The purpose of this section is to give a realistic example of the use of discrete-time cohort decomposition for cost-effectiveness analysis. For this purpose, we replicate a study by Schousboe and colleagues,<sup>9</sup> who conducted a cost-effectiveness analysis of alendronate therapy for fracture risk in osteopenic postmenopausal women. Schousboe and colleagues considered 5 different types of fractures: hip fracture, clinical vertebral (Cv) fracture, radiographic vertebral (Rv) fracture, distal forearm (Df) fracture, and other fracture. As we have already mentioned, these authors did not conduct a factored analysis, but it is natural to model the situation in this way, with one factor for each fracture type. Cohort decomposition requires factors to be independent, which in this case means that there is no dependence of the rate of one type of fracture on whether fractures of other types have occurred. This is not completely realistic but is in fact exactly the assumption that Schousboe and colleagues invoked, as there were no available data to estimate this dependence. We therefore retain this assumption for our analysis.

We included a background mortality factor along with 1 factor for each of the 5 fracture types. With each fracture factor having 2 nondeath states (no fracture and post fracture) there are  $1 \times 2^5$  combinations of nondeath states plus 1 overall death state, for a total of 33 combination states and therefore 33 columns in a spreadsheet should one wish to conduct a naive cohort analysis. Cohort decomposition, on the other hand, requires 3 columns for hip fracture (no fracture, post fracture, and death), 2

columns each for each of the remaining 4 fracture types (which do not involve death), and 2 columns for background mortality, for a total of 13 columns, more than a 50% reduction in column requirements. Moreover, as we have pointed out already, there is an accompanying simplicity in the cohort analyses themselves, all of which involve transitions among only 2 or 3 states.

Figure 7 illustrates both a continuous-time and a discrete-time transition diagram for the hip fracture factor. In Figure 7a, there is an age-dependent transition rate  $\lambda_{Hip}(t)$  from no fracture to hip fracture, followed by an immediate age-dependent mortality probability  $p_{MHip}(t)$  equal to 1.375 times the yearly background mortality probability. In the state post-hip fracture, the incidence rate of a subsequent hip fracture is higher by a relative risk of  $rr_{Hip} = 1.7$ . The dashed line around the second hip fracture state indicates that it is the same state as the first hip fracture. So, for example, a second hip fracture results in death with the same probability  $p_{MHip}(t)$ .

The discrete-time transition diagram of Figure 7b is obtained from the diagram of Figure 7a using Poisson decomposition (see Hazen<sup>16</sup>). The transition probabilities in this diagram would be obtained by converting the corresponding transition rates in the standard way.<sup>1,4</sup>

For the 4 other fracture factors, the transition diagrams are identical and shown in Figure 8. In Figure 8a, the incidence rate  $\lambda_F(t)$  from no fracture to fracture is age dependent. After the first fracture, individuals go to the state of post fracture, where the incidence rate is enhanced by a relative risk of  $rr_F$ . The discrete-time transition diagram in Figure 8b is obtained once again by Poisson decomposition.

Therapy with alendronate lasts for 5 years, during which time the incidence of vertebral fractures is reduced by a factor of  $RR_v = 0.54$  for women with T scores of  $-2.0$ , which is the group we considered in our replication (Schousboe and colleagues<sup>9</sup> also considered T scores of  $-1.5$  and  $-2.4$ ). After that, a linear, gradual offset of fracture reduction benefit over the subsequent 5 years is assumed.

Parameter values are listed in Tables 1 and 2. For the fracture risk functions  $\lambda_F(t)$ , see the appendix in Schousboe and colleagues.<sup>9</sup> On the basis of these, we used the cohort decomposition methods in the "Discrete-Time Cohort Decomposition" section to obtain expected costs and QALYs with and without alendronate for a hypothetical cohort of ten thousand 55-year-old osteopenic postmenopausal white women with a femoral T score of  $-2.0$ .

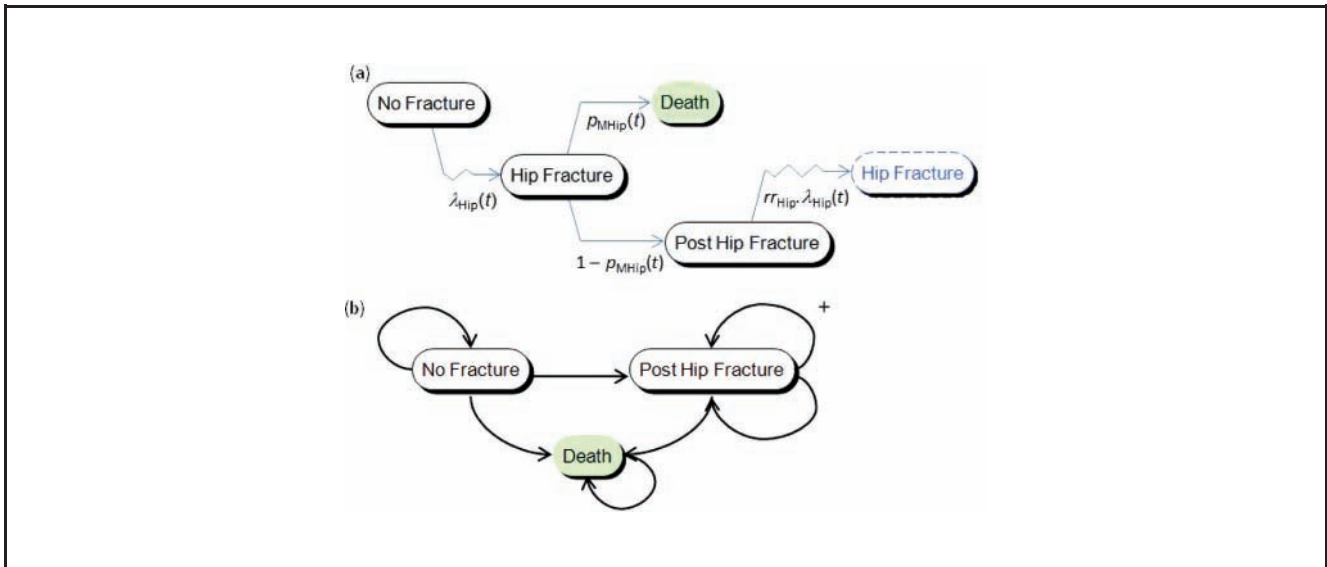


Figure 7 (a) Continuous-time transition diagram for the factor of hip fracture. Here the dashed line around the hip fracture state denotes that it is the same state as the previous hip fracture state. (b) Corresponding discrete-time transition diagram obtained by Poisson decomposition of (a) and converting rates to probabilities. Here the arrow from post-hip fracture labeled (+) denotes another fracture occurrence, whereas the accompanying unlabeled arrow denotes remaining in post-hip fracture without fracture for another cycle.

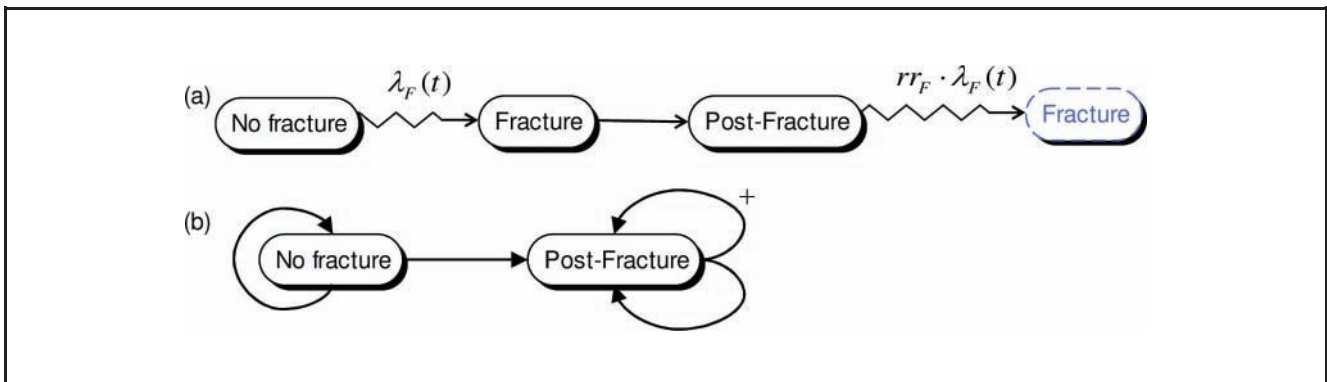


Figure 8 (a) Continuous-time transition diagram for other fractures. (b) Corresponding discrete-time transition diagram. The label (+) denotes another fracture occurrence, whereas the accompanying unlabeled arrow denotes remaining in post fracture for 1 cycle without fracture.

The only quality deviations from full health in this model were quality tolls due to fractures. Therefore, the overall QALY coefficients  $v(x)$  were equal to 1 for all composite states  $x = (x_1, x_2, x_3, x_4, x_5, x_6)$ , and the multiplicative restriction (4) holds automatically with  $v_i(x_i) = 1$  for all  $x_i$  and all  $i$ . For this reason, there is no numerical distinction between absolute and proportional QALY tolls. Also, overall ongoing QALYs are equal to life years in our results.

The results of our cohort analysis are shown in Table 3. Consistent with the analysis by Schousboe and colleagues,<sup>9</sup> the benefit of alendronate is

small—only 1/30 of a year, or about 12 days of life. However, the incremental cost is also small, resulting in an incremental cost-effectiveness ratio (ICER) value of \$110,068/life year.

### DISCUSSION

We have introduced cohort decomposition for Markov cost-effectiveness models, a technique that, when applicable, can decrease computational effort, facilitate and simplify model formulation, reduce

**Table 1** Relative Risks and Discount Rates for Our Replication of Schousboe and Colleagues' Analysis<sup>9</sup> of Alendronate for Fracture Risk in Osteopenic Postmenopausal Women

Relative Risks		
RRNs	For nonspinal fractures under alendronate	1.0
RRV	For vertebral fractures under alendronate	0.54
rrHip	For subsequent hip fracture	1.7
rrCv	For subsequent Cv fracture	4.0
rrRv	For subsequent Rv fracture	1.0
rrDf	For subsequent Df fracture	2.05
rrOther	For subsequent other fracture	1
rrMuHip	For background mortality in the first year subsequent to hip fracture	1.375
Discount Rates		
rCost	Costs	0.03/yr
rQALY	QALYs	0.03/yr

For the values of age-dependent fracture risks, see Schousboe and colleagues.<sup>9</sup> Cv, clinical vertebral; Rv, radiographic vertebral; Df, distal forearm; QALY, quality-adjusted life year.

**Table 2** Cost and QALY Parameters for Our Replication of Schousboe and Colleagues' Analysis<sup>9</sup> of Alendronate for Fracture Risk in Osteopenic Postmenopausal Women

Ongoing QALY			
vNo	No fracture state		1.00
vHip	Post-hip fracture state		0.81
vCv	Post-Cv fracture state		0.90
vRv	Post-Rv fracture state		1.00
vDf	Post-Df fracture state		1.00
vOther	Post-other fracture state		0.97
vAlive	Alive in background mortality		0.84
QALY Tolls			
qHip	Hip fracture		0.012
qCv	Cv fracture		0.214
qRv	Rv fracture		0.456
qDf	Df fracture		0.023
qOther	Other fracture		0.071
Ongoing Costs			
cAlen	Alendronate	\$842/year	
cMoHip	Care after hip fracture	\$6295/year	
cPhyVisit	Physician visit	\$52/year	
Cost Tolls			
		Direct	Indirect
cInitHip	Hip fracture	\$25,027	Age dependent <sup>a</sup>
cInitCv	Cv fracture	\$6702	Age-dependent <sup>a</sup>
cInitDf	Df fracture	\$3276	Age-dependent <sup>a</sup>
cInitOther	Other fracture	\$5561	Age-dependent <sup>a</sup>

Cv, clinical vertebral; Rv, radiographic vertebral; Df, distal forearm; QALY, quality-adjusted life year.

<sup>a</sup>See Appendix Table 4 in Schousboe and colleagues.<sup>9</sup>

model complexity, facilitate graphical depiction of models, and thereby open model structure to peer inspection and critique.

It is true that computational savings may not be a crucial issue for many of the smaller models that appear in the literature, and that, if necessary, software platforms other than spreadsheets may suffice

if computation time is a problem. Nevertheless, as we have discussed, regardless of the platform, computational savings can be considerable for larger models. This may be important in particular for computation of value of sample information or of value of information on subsets of parameters,<sup>17,18</sup> which can be arduous even for smaller models.

**Table 3** Results for Our Replication of Schousboe and Colleagues' Analysis<sup>9</sup> of Alendronate for Fracture Risk in Osteopenic Postmenopausal Women

	No Alendronate	Alendronate		No Alendronate	Alendronate
<b>Costs</b>	<b>\$18,342</b>	<b>\$22,008</b>	<b>QALYs</b>	<b>14.9787</b>	<b>15.0120</b>
<b>Cost Tolls</b>	<b>\$9,960</b>	<b>\$9,763</b>	<b>QALY Tolls</b>	<b>0.2004</b>	<b>0.1854</b>
<b>Hip fracture</b>	<b>\$4,776</b>	<b>\$4,776</b>	<b>Hip fracture</b>	<b>0.0016</b>	<b>0.0016</b>
First	\$3,409	\$3,409	First	0.0012	0.0012
Subsequent	\$1,367	\$1,367	Subsequent	0.0004	0.0004
<b>Cv fracture</b>	<b>\$1,912</b>	<b>\$1,714</b>	<b>Cv fracture</b>	<b>0.0464</b>	<b>0.0422</b>
First	\$1,155	\$1,034	First	0.0277	0.0253
Subsequent	\$757	\$680	Subsequent	0.0187	0.0168
<b>Df fracture</b>	<b>\$586</b>	<b>\$586</b>	<b>Rv fracture</b>	<b>0.1243</b>	<b>0.1134</b>
First	\$491	\$491	First	0.0941	0.0862
Subsequent	\$95	\$95	Subsequent	0.0302	0.0273
<b>Other fracture</b>	<b>\$2,687</b>	<b>\$2,687</b>	<b>Df fracture</b>	<b>0.0028</b>	<b>0.0028</b>
First	\$1,845	\$1,845	First	0.0023	0.0023
Subsequent	\$842	\$842	Subsequent	0.0005	0.0005
<b>Ongoing Costs</b>	<b>\$8,382</b>	<b>\$12,245</b>	<b>Other fracture</b>	<b>0.0253</b>	<b>0.0254</b>
<b>Hip fracture</b>	<b>\$7,411</b>	<b>\$7,411</b>	First	0.0171	0.0171
<b>Physician Visit</b>	<b>\$971</b>	<b>\$971</b>	Subsequent	0.0082	0.0082
<b>Alendronate</b>	<b>\$0</b>	<b>\$3,863</b>	<b>Ongoing QALYs</b>	<b>15.1791</b>	<b>15.1973</b>
<b>ΔC =</b>	<b>\$3,665</b>			<b>ΔQ =</b>	<b>0.0333</b>
				<b>ΔC/ΔQ =</b>	<b>\$110.068</b>

The table gives lifetime costs and quality-adjusted life years (QALYs) for a cohort of 55-year-old osteopenic postmenopausal white women with a femoral T score of -2.0 with and without alendronate treatment. Cv, clinical vertebral; Rv, radiographic vertebral; Df, distal forearm.

Moreover, advantages other than computational savings are just as important, in our view. Compared to the overall model, individual factors typically have only a few states and transitions, resulting in considerably simpler model formulation. State proliferation is an issue for any large model, as we have pointed out earlier. Instead of restricting transitions to a subset of states or abandoning Markov cohort analysis for patient-level simulation, a modeler can use a factored formulation to constructively accommodate state proliferation. The popular software program TreeAge accommodates factoring via so-called *clones* and *clone masters*.

Because graphical depiction is possible, model presentation is thereby considerably simpler as well, and this simplicity opens the model structure to peer inspection and critique. In our experience, it is rare for a published Markov cost-effectiveness analysis of any complexity to even attempt to give a graphical picture of model structure—only the simplest analyses do so. We suspect it is correspondingly rare for reviewers to critically examine model assumptions when there is no graphical summary of those assumptions. The techniques of this article provide a graphical vehicle to yield such a summary. These graphical techniques can be employed even if the probabilistic independence assumptions required for cohort decomposition do not hold.

Although the multiplicative assumption (4) needed for cohort decomposition of ongoing QALYs may appear unduly restrictive, in fact it often holds by default. For instance, as we noted above in our reanalysis of Schousboe and colleagues,<sup>9</sup> (4) will hold when all quality coefficients are equal to 1 (i.e., when one-time tolls fully capture all decrements in quality from full health). It will also hold whenever ongoing quality decrements are restricted to exactly one of the factors, say factor *i*, in which case one can set  $v_i(x_i)$  to the desired quality level and take  $v_j(x_j) = 1$  for all  $j \neq i$ . The latter is a common occurrence in our experience.

### APPENDIX Cohort Decomposition Formulas

We derive here the continuous-time formulas for cohort decomposition. To obtain discrete-time formulas, replace  $\int \dots e^{-rt} dt$  with  $\sum \dots (1+r)^{-t} \Delta t$  and transition rates  $q_{xy}(t)$  by transition probabilities  $p_{xy}(t)$ . The formulas produce costs and QALYs at the per person level, for which the following terminology is useful. Let  $\mathbf{1}_x(t)$  be the indicator function for an individual occupying state *x* at time *t*, that is,

$$\mathbf{1}_x(t) = \begin{cases} 1 & \text{if state } x \text{ is occupied at time } t \\ 0 & \text{if not.} \end{cases}$$

If  $P_x(t)$  is the probability that an individual occupies state  $x$  at time  $t$ , then

$$P_x(t) = P(\mathbf{1}_x(t) = 1) = E[\mathbf{1}_x(t)].$$

In a cohort analysis with  $N$  individuals, the relationship of this probability to the expected number  $n_x(t)$  in state  $x$  at time  $t$ , computed via the cohort analysis formulas (1), is

$$P_x(t) = \frac{1}{N} n_x(t).$$

Also, let  $\mathbf{1}_j(t)$  be the indicator function of factor  $j$  being in a nondeath state at time  $t$ , and given a specific factor  $i$  under consideration, let  $\mathbf{1}_O(t) = \prod_{j \neq i} \mathbf{1}_j(t)$  be the indicator

function of all other factors occupying nondeath states at time  $t$ .

Throughout the following derivations, we shall assume that factors are probabilistically independent. If  $S_j(t)$  is the probability that factor  $j$  is in a nondeath state at time  $t$  (the survival function factor  $j$  at time  $t$ ), then we have

$$E[\mathbf{1}_O(t)] = E\left[\prod_{j \neq i} \mathbf{1}_j(t)\right] = \prod_{j \neq i} E[\mathbf{1}_j(t)] = \prod_{j \neq i} S_j(t).$$

### Ongoing QALY

Let  $x = (x_1, \dots, x_n)$  be the overall state vector, where  $x_i$  is the corresponding state in factor  $i$ . Then the indicator function  $\mathbf{1}_x(t)$  for the overall state being  $x$  at time  $t$  is equal to the product over all factors of the individual indicator functions  $\mathbf{1}_{x_i}(t)$ , that is,

$$\mathbf{1}_x(t) = \prod_i \mathbf{1}_{x_i}(t).$$

We treat here the common case in which the overall quality coefficient  $v(x)$  is the product across factors of quality coefficients  $v_i(x_i)$ , that is,

$$v(x) = \prod_i v_i(x_i). \quad (\text{A1})$$

The quantity we desire is the expected ongoing QALY per person, where

$$\text{Ongoing QALY/person} = \int_0^{\infty} e^{-rt} QR(t) \cdot dt. \quad (\text{A2})$$

Here  $QR(t)$  is the ongoing QALY accrual rate per person at time  $t$ , given by

$$QR(t) = \sum_x v(x) \mathbf{1}_x(t).$$

Let  $QR_i(t)$  be the individual QALYs accrual rate for factor  $i$  at time  $t$ ,

$$QR_i(t) = \sum_{x_i} v_i(x_i) \mathbf{1}_{x_i}(t). \quad (\text{A3})$$

Due to (A1), we have the decomposition

$$\begin{aligned} \sum_x v(x) \mathbf{1}_x(t) &= \sum_x \prod_i v_i(x_i) \mathbf{1}_{x_i}(t) \\ &= \prod_i \sum_{x_i} v_i(x_i) \mathbf{1}_{x_i}(t), \end{aligned}$$

that is,

$$QR(t) = \prod_i QR_i(t). \quad (\text{A4})$$

Take expectations in (A2), (A4), and (A3) to get the following cohort decomposition formulas:

$$E[\text{Ongoing QALY/person}] = \int_0^{\infty} e^{-rt} E[QR(t)] \cdot dt,$$

$$E[QR(t)] = \prod_i E[QR_i(t)],$$

$$E[QR_i(t)] = \sum_{x_i} v_i(x_i) P_{x_i}(t) = \frac{1}{N_i} \sum_{x_i} v_i(x_i) n_{x_i}(t),$$

where  $N_i$  is the cohort size in factor  $i$ .

### Ongoing Cost

Suppose there is a cost rate  $c_x$  per person for each unit of time in state  $x$  of factor  $i$  when other factors are in non-death states. Then ignoring other factors, the ongoing individual cost rate for factor  $i$  at time  $t$  is

$$CR_i(t) = \sum_x c_x \mathbf{1}_x(t).$$

Accounting for the fact that other factors must be in non-death states for costs in factor  $i$  to be incurred, we have

$$\text{Ongoing cost/person for factor } i = \int_0^{\infty} e^{-rt} CR_i(t) \mathbf{1}_O(t) dt.$$

Taking expectations in the last 2 equations gives the following cohort decomposition formulas:

E[Ongoing cost/person for factor  $i$ ]

$$= \int_0^\infty e^{-rt} E[CR_i(t)] \prod_{j \neq i} S_j(t) dt,$$

$$E[CR_i(t)] = E\left[\sum_x c_x \mathbf{1}_x(t)\right] = \sum_x c_x P_x(t) = \frac{1}{N} \sum_x c_x n_x(t).$$

**Cost Tolls**

Suppose there is a cost toll  $c_{xy}$  for transition  $x \rightarrow y$  in factor  $i$ . Then the cost toll rate per person for factor  $i$  at time  $t$  satisfies

$$CTR_i(t) \cdot dt = \sum_x \sum_y c_{xy} dN_{xy}(t).$$

Here  $dN_{xy}(t)$  is the number of transitions  $x \rightarrow y$  in  $[t, t + dt)$ . If the transition rate from  $x$  to  $y$  at time  $t$  is  $q_{xy}(t)$ , then  $E[dN_{xy}(t)] = P_x(t)q_{xy}(t)dt$ . Accounting for the fact that cost tolls are incurred only if no other factors are in death states, we have

$$\text{Cost toll/person for factor } i = \int_0^\infty e^{-rt} (CTR_i(t)) \cdot \mathbf{1}_O(t) dt.$$

Take expectations of the last 2 equations to get the following cohort decomposition formulas:

E[Cost toll/person for factor  $i$ ]

$$= \int_0^\infty e^{-rt} E[CTR_i(t)] \cdot \prod_{j \neq i} S_j(t) \cdot dt,$$

$$E[CTR_i(t)] = \sum_x P_x(t) \sum_y c_{xy} q_{xy}(t) = \frac{1}{N_i} \sum_x n_x(t) \sum_y c_{xy} q_{xy}(t).$$

**QALY Tolls**

Suppose there is a QALY toll  $\delta_{xy}$  for transition  $x \rightarrow y$  factor  $i$  when the remaining factors are alive. Then the QALY toll rate per person for factor  $i$  at time  $t$  is

$$QTR_i(t) \cdot dt = \sum_x \sum_y \delta_{xy} dN_{xy}(t).$$

Accounting for the fact that cost tolls are incurred only if no other factors are in death states, we have

$$\text{QALY toll/person for factor } i = \int_0^\infty e^{-rt} QTR_i(t) \cdot \mathbf{1}_O(t) \cdot dt.$$

Take expectations of the last 2 equations to get the following cohort decomposition formulas:

E[QALY toll/person for factor  $i$ ]

$$= \int_0^\infty e^{-rt} E[QTR_i(t)] \cdot \prod_{j \neq i} S_j(t) \cdot dt,$$

$$E[QTR_i(t)] = \sum_x P_x(t) \sum_y \delta_{xy} q_{xy}(t) = \frac{1}{N_i} \sum_x n_x(t) \sum_y \delta_{xy} q_{xy}(t).$$

In the proportional toll situation mentioned in the text, it may be more natural to assume that the overall toll produced by a transition  $x_i \rightarrow y_i$  in factor  $i$  is a toll  $\delta_{x_i y_i}$  in factor  $i$  times the product  $\prod_{j \neq i} v_j(x_j)$  of QALY coefficients in other factors in states  $x_j$ . In this case, an alternate formula applies:

QALY toll/person for factor  $i$

$$\begin{aligned} &= \int_0^\infty e^{-rt} \sum_{x_i} P_{x_i}(t) \sum_{y_i} \delta_{x_i y_i} q_{x_i y_i}(t) \cdot \prod_{j \neq i} \sum_{x_j} v_j(x_j) \mathbf{1}_{x_j} \cdot dt \\ &= \int_0^\infty e^{-rt} QTR_i(t) \cdot \prod_{j \neq i} QR_j(t) \cdot dt. \end{aligned}$$

Taking expectations then yields the alternate cohort decomposition formula:

E[QALY toll/person for factor  $i$ ]

$$= \int_0^\infty e^{-rt} E[QTR_i(t)] \cdot \prod_{j \neq i} E[QR_j(t)] \cdot dt.$$

**REFERENCES**

1. Hunink MGM, Glasziou P, Siegel JE, et al. Decision Making in Health and Medicine: Integrating Evidence and Values. Cambridge, UK: Cambridge University Press; 2001.
2. Beck J, Pauker SG. The Markov process in medical prognosis. Med Decis Making. 1983;3:419–53.
3. Beck J, Sonnenberg FA. Markov models in medical decision making: a practical guide. Med Decis Making. 1993;13: 322–38.
4. Briggs A SM, Claxton K. Decision Modeling for Health Economic Evaluation. Oxford, UK: Oxford University Press; 2006.
5. Col NF, Goldberg RJ, Orr RK, Erban JK, Fortin JM, Chlebowski RT. Survival impact of tamoxifen use for breast cancer risk



- reduction: projections from a patient-specific Markov model. *Med Decis Making*. 2002;22:386–93.
6. Anderson K, Jacobson JS, Heitjan DF, et al. Cost-effectiveness of preventive strategies for women with a BRCA1 or a BRCA2 mutation. *Ann Intern Med*. 2006;144:397–406.
7. Tsevat J, Eckman MH, McNutt RA, Pauker SG. Warfarin for dilated cardiomyopathy: a bloody tough pill to swallow? *Med Decis Making*. 1989;9:162–9.
8. Hazen GB. Factored stochastic trees: a tool for solving complex temporal medical decision models. *Med Decis Making*. 1993;13:227–36.
9. Schousboe JT, Nyman JA, Kane RL, Ensrud KE. Cost-effectiveness of alendronate therapy for osteopenic postmenopausal women. *Ann Intern Med*. 2005;142:734–41.
10. Roach PJ, Fleming C, Hagen MD, Pauker SG. Prostatic cancer in a patient with asymptomatic HIV infection: are some lives more equal than others? *Med Decis Making*. 1988;8:132–44.
11. Sonnenberg FA, Pauker SG. Elective pericardiectomy for tuberculous pericarditis: should the snappers be snipped? *Med Decis Making*. 1986;6:110–23.
12. Plante D, Piccirillo J, Sofferman R. Decision analysis of treatment options in pyriform sinus carcinoma. *Med Decis Making*. 1987;7:74–83.
13. Fleming C, Wong J, Moskowitz A, Pauker S. A peripartum neurologic event: shooting from the hip. *Med Decis Making*. 1988;8:55–71.
14. Eckman M, Singh S, Erban J, Kao G. Testing for factor V Leiden in patients with pulmonary or venous thromboembolism: a cost-effectiveness analysis. *Med Decis Making*. 2002;22:108–24.
15. National Cancer Institute. Surveillance, Epidemiology, and End Results (SEER) Program. Bethesda, MD: National Cancer Institute; 2005.
16. Hazen GB. Stochastic trees: a new technique for temporal medical decision modeling. *Med Decis Making*. 1992;12:163–78.
17. Brennan A, Kharroubi S, O'Hagan A, Chilcott J. Calculating partial expected value of perfect information via Monte Carlo sampling algorithms. *Med Decis Making*. 2007;27:448–70.
18. Brennan A, Kharroubi SA. Efficient computation of partial expected value of sample information using Bayesian approximation. *J Health Econ*. 2007;26:122–48.