WORK IN PROGRESS: PLEASE DO NOT QUOTE WITHOUT AUTHORS' PERMISSION

An empirical comparison of a Markov process and a discrete event simulation for modelling the cost-effectiveness of alternative adjuvant therapies for early breast cancer.

Jon Karnon,

Health Economics Research Group, Brunel University, Uxbridge, UB8 3PH

Paper presented to the Health Economists' Study Group Conference,

University of Oxford, January 2001

Introduction

Modelling, as a decision analytic tool, is a key appliance in the health economists' toolbox[Buxton et al, 1997]. The critical evaluation of most aspects of the modelling process is under developed, but this problem appears to be particularly acute with respect to the appropriate choice of decision modelling technique. This situation is partially explained by the limited range of decision analytic modelling methods employed by health economists. A recent HESG paper reviewed modelling economic evaluations published in 1997[Barton et al, 2000]. Of the 119 papers reviewed it was reported that 76 (64%) employed decision trees, 43 (36%) used Markov processes and only 2 (2%) reported the results from discrete event simulation (DES) models[1].

The characteristics of decision trees and Markov processes are very different and the choice between the two techniques in alternative treatment settings is relatively straightforward. With the introduction of DES to the field of economic evaluation in health care the issue of choosing the appropriate technique could become an important decision in the initial stages of modelling projects. The issue of choosing the correct modelling technique has been referred to in the health economics literature[Sonnenberg et al, 1994;Chaussalet et al, 1999], but the consequences of the choice have not been fully explored.

Both DES models and Markov processes are forms of simulation, but DES allows more complicated representations of the system being modelled[Hillier and Lieberman, 1995]pg901. Within a DES model, patients move through the model, experiencing events at any discrete time period after the previous event. The analysis of the model is triggered by the occurrence of an event, at which point the model asks what and when is the next event for this patient, rather than a Markov process, which asks what events are occurring at regular intervals.

This paper aims to investigate the effect of the choice of employing a Markov process or a DES model to represent patient pathways over extended time horizons on the process and output of modelling projects.

---

[1] The total is 121 because two studies reported results from both a decision tree and a Markov chain.

The author hypothesised that alternative modelling techniques may be compared on the basis of two broad criteria – flexibility and analytic input. Each covers various aspects:

- Flexibility describes how well the model adapts to different forms of input data and interrelationships between parameters. This criterion reflects how closely a particular modelling technique allows the reality of patient pathways to be modelled.
- Analytic input relates to the level of complexity in building the model in terms of the level of expertise and the amount of time required. This criterion is also influenced by the ease with which modifications to the model can be made, both to the actual structure of the model and to the incorporation of alternative formats of model inputs.

The economic evaluation of tamoxifen alone versus tamoxifen and chemotherapy in node positive, postmenopausal women is used as the basis for the comparison of the modelling techniques. Breast cancer is the commonest female cancer in the United Kingdom with around 33,000 newly diagnosed cases and 15,000 deaths from the disease each year[Cancer Research Campaign, 1998]. Following diagnosis of breast cancer in the breast and/or axilla a proportion of patients will be 'cured' by local treatment. However, there is a risk of micrometastatic disease, present in distant areas, causing systemic relapse. The aim of adjuvant therapy is to destroy this subclinical disease[Fox, 1991]. The main objective of adjuvant therapy for breast cancer is to prolong survival while maintaining a high quality of life[Glick, 1991].

The comparison of the modelling techniques is also presented in the context of a stochastic evaluation. A number of papers have described the use of probabilistic sensitivity analysis in economic evaluation[Felli and Hazen, 1998;Lord and Asante, 1999;Pasta et al, 1999]. Alternatively labelled stochastic cost-effectiveness analyses[Briggs, 1999], the value of the input parameters within such models are described as probability distributions. The objective of a stochastic analysis is to obtain a distribution of each of the model's outputs that are informed by randomly sampled sets of input parameter values from the specified probability distributions.

Methods

The methods are described in two sections covering the sequential development of the alternative decision models - building and populating.

Building the models

The structure of the models was based on the inclusion of events that have a significant impact on the outcomes of the model. Information on the relevant events included in the decision model was sought from health professionals as well as a preparatory review of the breast cancer literature. In addition, basic information on the impact and progression of breast cancer was obtained from the Internet. Details of the adopted model structures are presented in the following sections.

The specification of a minimum time period of advancement was required for the DES model, as well as a cycle length for the Markov process. Though a case could be made for differential timing on the basis of model running time, the author decided that patient pathways could be most accurately represented in both models using a time period of one month. A similar time horizon for both models was also adopted with patients living to a maximum of 100 years.

Separate descriptions of the process of building the two models are presented in the following sections. For each model, the description of the actual modelling of the health states is followed by the necessary structures for controlling the model inputs and outputs.

The Markov process

The Markov process was built in Excel using the Crystal Ball 2000 add-in, which is a risk analysis programme that is "easy to learn and easy to use"[Crystal Ball, 2000]. Crystal Ball uses Monte Carlo simulation to facilitate stochastic analyses.

Modelling the health states

The model structure, presented in figure 1, shows the possible transitions within the model moving from period 't' to 't+1' at different phases of the model. All patients start the model in the health state disease free interval (DFI). During the first six months patients in the DFI state may experience toxic effects caused by the adjuvant therapies, a relapse in one of the four specified sites of relapse – locoregional, soft tissue, bone or visceral, or they may die from a cause other than breast cancer. Within the first six months patients may experience a relapse or die from other causes from the toxicity states. At the end of the first six months the patients in the toxicity states are assumed to return to the original DFI state.

Within the Markov process, the author assumed that different forms of toxicity could not be experienced simultaneously. There were for two reasons for this assumption. Firstly, no data describing the proportion of patients experiencing joint toxicities were available. Secondly, four extra health states would be required to represent the occurrence of the joint toxicities.

Other than the probability of experiencing toxicity, patients in the DFI state are subject to monthly probabilities of moving to a relapse state or straight to the dead state (with no evidence of disease).

Patients experiencing locoregional relapses are assumed to remain in that state for a maximum of one month, following which time they enter a period of remission, they experience a more severe metastatic relapse, or they die from other causes. The majority of the data identified in the literature review described the time in remission in the form of a survival curve. Due to the Markovian assumption, which states that future pathways are dependent only on the status of the current state and the overall time spent within the model, it was necessary to convert the data represented by the survival curves into a constant probability of experiencing a further event.

From remission, patients can experience metastases or die from other causes. Patients in the metastatic states patients can only progress to death. The need for further data conversions arose with respect to the data describing survival times from metastases.

Figure 1 Representation of the possible patient pathways within the ABC Markov process

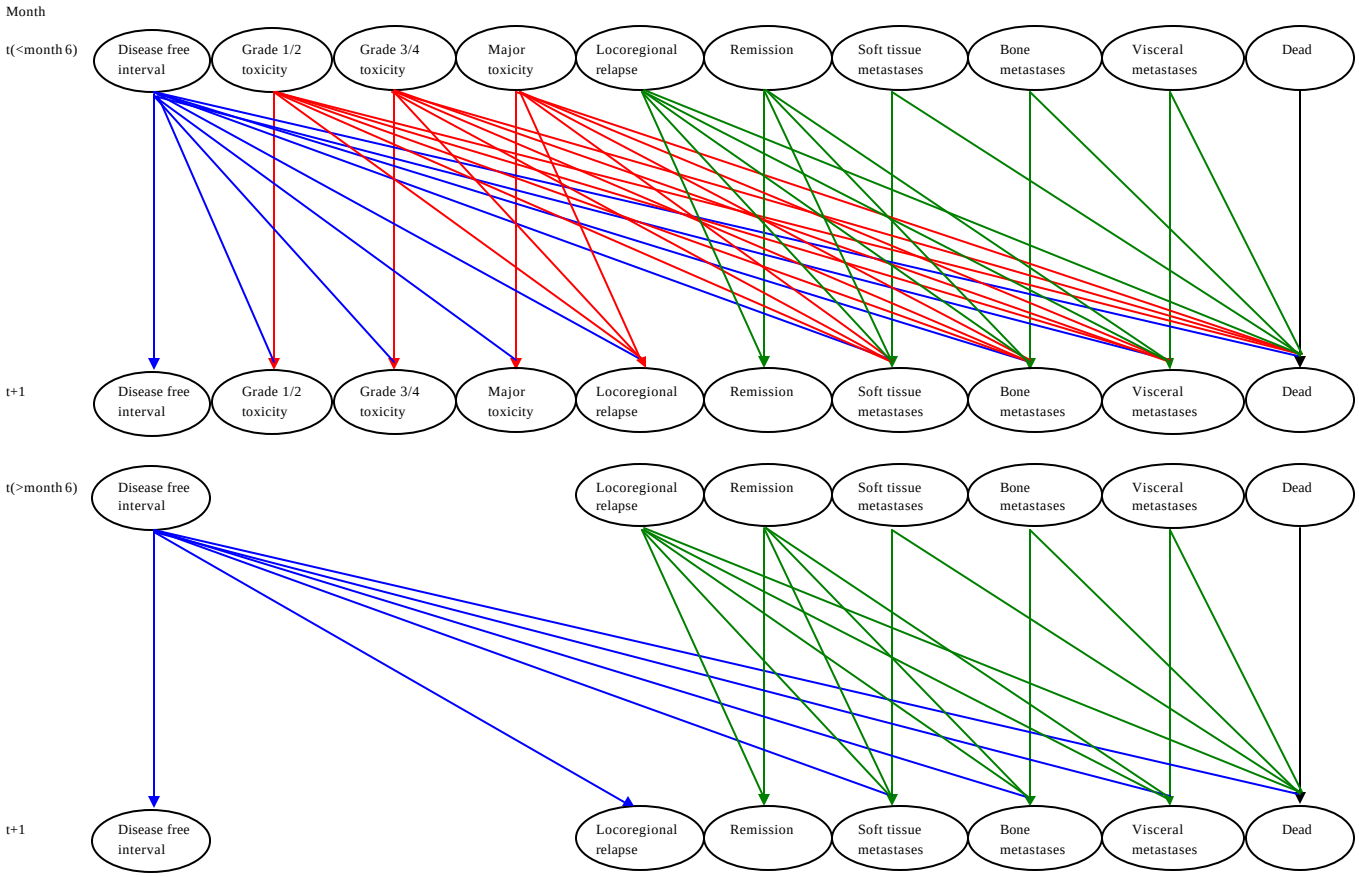
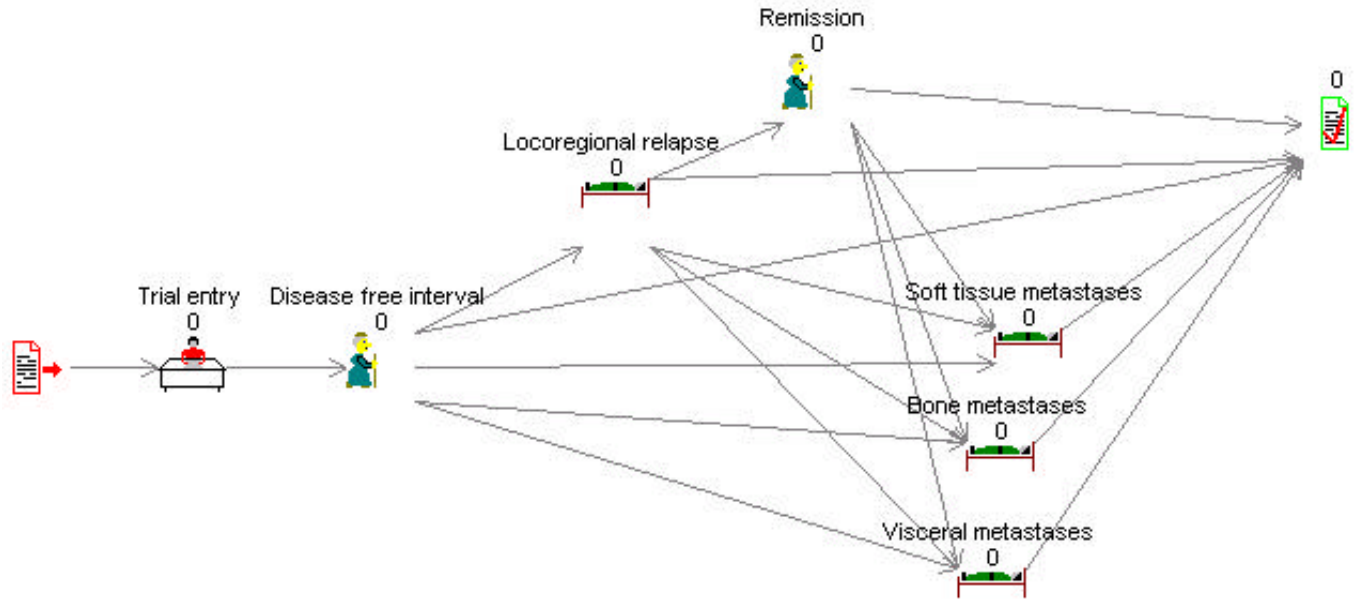


Figure 2 Representation of the possible patient pathways within the ABC DES model



Most of the data in the literature were in the format of a median survival time from diagnosis with metastases in the alternative sites, which required conversion to a constant monthly probability of dying.

Controlling inputs and outputs

The Markov process ran on one spreadsheet, wherein each health state was assigned a column and each successive cycle (month) was represented by a row. From this allocation of columns and rows the basic operation of the Markov process involved the description of the proportion of patients placed in each health state during each time cycle of the model. Separate spreadsheets were established to hold the relevant data on clinical parameters, costs and utility values. Within each of the parameter spreadsheets the relevant probability distributions were established using the facilities within the Crystal Ball add-in[Crystal Ball, 2000].

Stochastic analysis comprises multiple deterministic analyses using alternative sets of parameter values that were randomly sampled from the defined probability distributions for each input parameter. Each deterministic analysis is defined as a 'run'.

A deterministic analysis (run) of a Markov process involves the following steps:

1. The proportion of patients in each state in each cycle are multiplied by the assigned monthly cost and utility values;
2. The respective costs and utility values are then summed for each cycle;
3. Separate discount factors are applied to the cost and utility values assembled for each month;
4. Aggregate values for the model outputs are calculated by summing the costs, lifeyears and utilities for all cycles.

10,000 alternative sets of input parameter values were sampled and Crystal Ball ran the necessary number of Monte Carlo simulations 'solving' the model for each run.

A full model was built and submitted for verification and validation in less than 2 weeks. Using a 700MHZ PC with a pentium II processor 10,000 runs were completed in approximately 1 hour.

The discrete event simulation model

The first task in building a DES model was to identify an appropriate software package that could handle the characteristics of an economic HTA decision model, and which did not require a high-level of programming skills. Simul8 software was chosen[Simul8, 2000]. The process of learning to use the software was undertaken on a separate project that employed a simpler model comparing a hospital at home scheme with inpatient care for elderly patients[Campbell et al, 2000]. The process was gradual and continued into the time spent building the ABC DES model.

Modelling the health states

The structure of the DES model is presented in figure 2, which is the visual interface of the simulation software. Relationships between states, activities within states, and the collection of data associated with each state were handled by programming code alongside the occurrence of the states. Code can be implemented as patients enter a state, when they are within a state, or as they leave a state. This will become clearer as the process of building the ABC model unfolds.

The model was built up in modules[Pidd, 1989]. The first, and main, module described the period of time spent in the state disease free interval (DFI). DFI is not the first state in the DES model because it is easier to assign the time spent in DFI prior to patients entering the state.

The most noticeable difference between the DES model and the Markov process is that fewer health states were described in the DES model. In the DES model, toxicity was modelled as an attribute of the health state DFI, rather than as individual states. Incorporating toxicity within DFI allowed each patient within the model to be subject to separate probabilities of experiencing each type of toxicity. The associated costs of each event experienced were attached to the patient. The QALYs associated with the

months spent in DFI was a function of the length of the different types of toxicity. For example, for a patient experiencing two types of toxicity, the utility value associated with the less severe form of toxicity would only be required if the lesser toxicity lasted longer than the more severe form.

The second module represented the passage of events from the point of locoregional relapse to either a metastatic relapse or death. The DES model also assumed that patients remain in the locoregional relapse state for exactly one month before moving onto the remission state. During the locoregional relapse state the relevant cost and utility value is attached to patients, and, analogous to the role of the trial entry state, the length of remission is determined using a similar process of looping through months until an event is experienced. The disease-free survival curves available to describe the time in remission were assimilated directly into the DES model up to 11 years post-diagnosis (end of reported follow-up). After 11 years the author assumed patients were subject to the mortality rate in the general population because the probability of experiencing a relapse was deemed to be insignificant.

The final module described the experience of patients within the three metastatic relapse states. As the only possible exit state from these states is death they are relatively easy to model. In the DES model, patients simply remain in the metastatic states for the sampled median survival time and then move into the dead state.

Controlling inputs and outputs

In the DES model, a single run (informed by a set of sampled input parameter values) followed 10,000 individual patients through the model. The mean model outputs for each group of 10,000 patients were the only outputs of interest for each run. The variation within the group of 10,000 patients represents first-order uncertainty, which does not provide useful information for the allocation of resources[Briggs, 2000]

.

The costs and QALYs associated with each patient were aggregated as patients left each state. This aspect of the programming code was the most complicated due to the need to discount the outputs. The process sectioned the time spent in each state into

the corresponding years that the patient has spent in the model. For example, if a patient entered a state in her 15th month in the model and left in her 33rd month then the first nine months were subject to a different discount factor than the second nine months.

2,500 runs of 10,000 patients, running the model day and night, required around three days. Around two months were required to build a model to the point at which the author had sufficient confidence to subject the model to the verification and validation process. Due to the learning curve effect the time building future models should be significantly reduced.

A table comparing the assumptions required by the alternative modelling techniques is presented in the appendix.

Populating the models

Separate reviews of the literature were undertaken for five components of the decision model - adjuvant therapies, treatment toxicity, timing of relapse or death with no evidence of cancer, types of relapse, and progression from relapse. Including all sources, at the time of the final analysis, a total of 343 full papers or documents had been reviewed. Additional primary sources for some of the cost parameters were used due to the small amount of relevant data in the literature.

For the stochastic analysis of the decision models, probability distributions were described to represent the identified data for each input parameter. The method adopted by the author applied theoretical considerations to the choice of probability distribution for different types of parameters. The process involved examining the characteristics of the different types of input parameters and assigning a particular type of probability distribution with properties that matched the input parameters. The distribution parameters for each probability distribution were informed by the available data for each input parameter.

The author identified four categories of parameters - proportions, survival times, costs and utility values - that are likely to found in a decision model. Survival times

describe the length of survival (or time to next event). An appropriate probability distribution chosen for parameter type:

1. Proportions describe the probability that a patient will experience an event, which is bounded by 0 and 1. Employing Bayesian methods for specifying prior distributions the beta distribution provides the most realistic representation of proportions[Iverson, 1984];
2. Survival times describe the length of survival (or time to next event), they are bounded by 0. The gamma distribution is bounded by zero and approximates the normal distribution at large samples. It is also extremely flexible, using a shape parameter to describe the available data accurately[Rice, 1995];
3. Cost parameters have been described by the lognormal distribution[Pasta et al, 1999;Fairclough, 1997], though the author has been advised that the gamma distribution provides a more flexible description of the sampling distribution of costs (personal communication: Andy Briggs);
4. Utility values portray similar properties to a proportion though 0 and 1 does not strictly bind them. The beta distribution is still advocated as a scale parameter can be fitted to the beta distribution to incorporate a larger range than 0 to 1.

The method of moments specifies formulae for estimating parameters for alternative probability distributions by finding expressions for them in their lowest order moments, then substituting sample moments into the expression[Rice, 1995]. The data identified from the literature review were combined with these formulae to estimate relevant parameters for the chosen probability distributions.

Results

The aggregate values for each of the model outputs derived from the two modelling techniques, from all three methods of specifying probability distributions, are presented in table 1.

Table 1 Comparison of the model outputs from the Markov process and DES model.

	Costs		QALYs		Lifeyears		ICERs*	
	Markov process	DES model	Markov process	DES model	Markov process	DES model	Markov process	DES model
Tamoxifen and chemotherapy	£9,146	£8,718	12.14	12.01	16.01	15.74	£3,334	£3,483

Tamoxifen alone	£7,115	£6,709	11.56	11.41	15.16	14.86

The mean cost, QALY, and lifeyear estimates do vary between the two modelling techniques, but in the same direction, ie. all three outputs are higher in the DES model. The cost differences are under £500 for both therapy options, and the differences in the QALY estimates are under 0.15 QALYs. These results lead to a very small divergence in the incremental cost-effectiveness ratios (ICERs) reported by the two models.

The credible intervals derived from the two models, reported in table 2, are also similar. Neither of the 2.5th percentiles exceeds £600. The lower limit of a confidence interval is unlikely to be of practical interest because the default resource decision is likely to be to provide tamoxifen alone. A one-sided hypothesis test at a 5% level of significance may then be deemed appropriate [Wakker and Klaassen, 1995], though even at the 95th percentile tamoxifen remains dominant on the basis of the intervals reported by both models.

Table 2 Credible intervals derived from the Markov process and the DES model

	2.5th percentile	95th percentile
Markov process	£452	tamoxifen dominates
DES model	£584	tamoxifen dominates

The cost and QALY differences between the two therapy options are plotted on the cost-effectiveness plane in figures 3a and 3b. These figures demonstrate the closeness in the ranges of cost-effectiveness between the two models. Though accounting for the reduced number of observations drawn from the DES model, it might be concluded that the observations derived from the DES model appear to be slightly more dispersed than those from the Markov process.

The cost-effectiveness acceptability (CEAc) curve provides a more flexible interpretation of cost-effectiveness presenting the probability that an intervention is cost-effective for all potential values of an additional unit of effect [Briggs and Fenn, 1998]. The CEAc curves presented in figure 4 show that the Markov process reports a slightly higher probability of tamoxifen and chemotherapy producing positive net benefits for all values of an additional QALY than the DES model. The values at

Figure 1a Observations of the cost and QALY differences plotted on the cost-effectiveness plane derived from the Markov process using theoretically specified input distributions

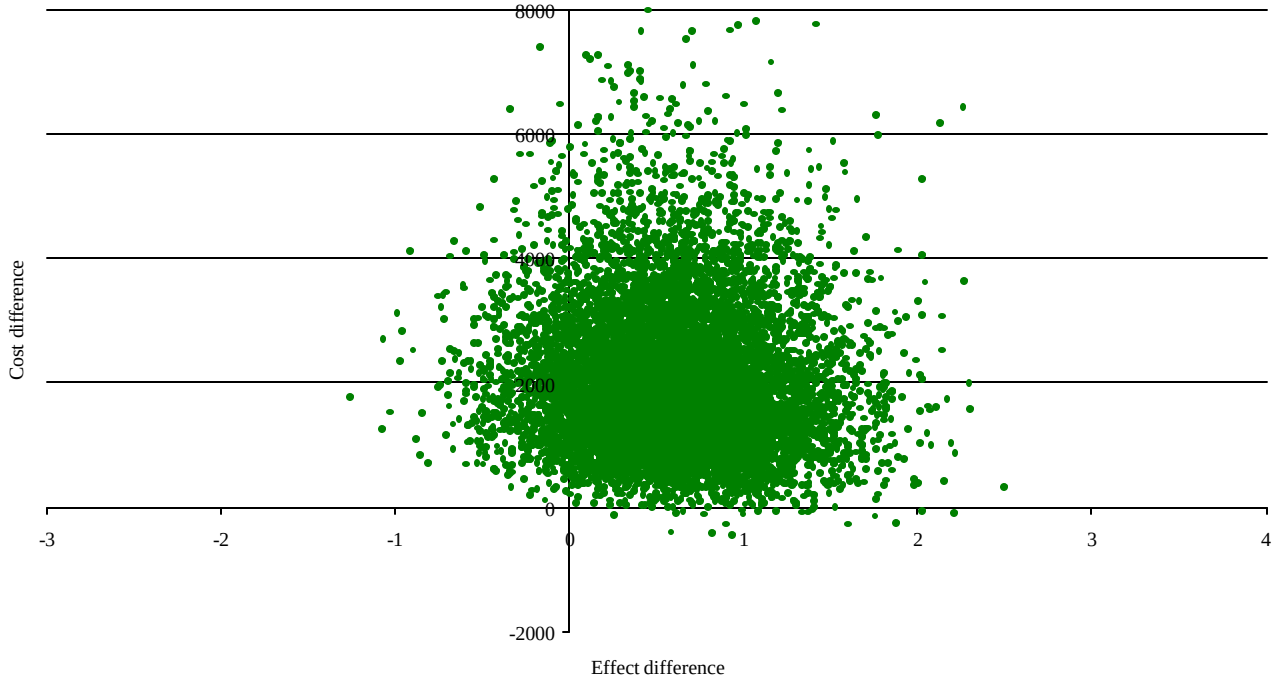


Figure 1b Observations of the cost and QALY differences plotted on the cost-effectiveness plane derived from the DES model using theoretically specified input distributions

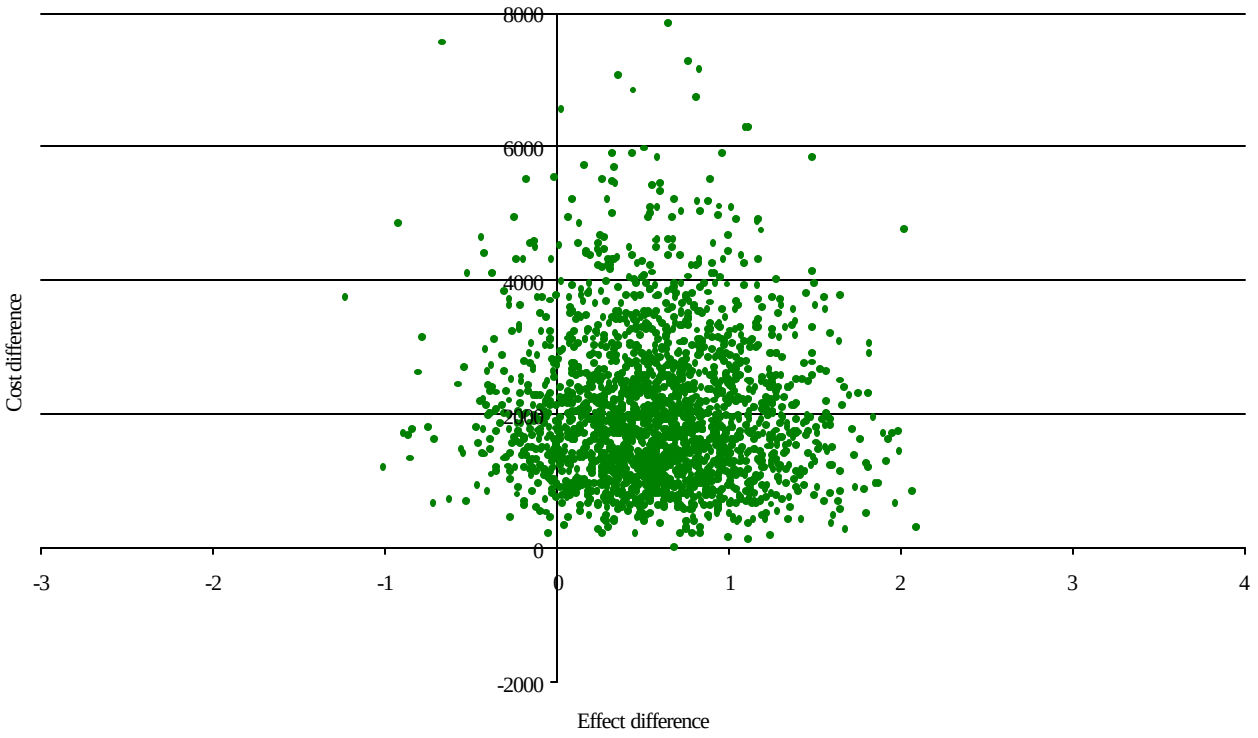
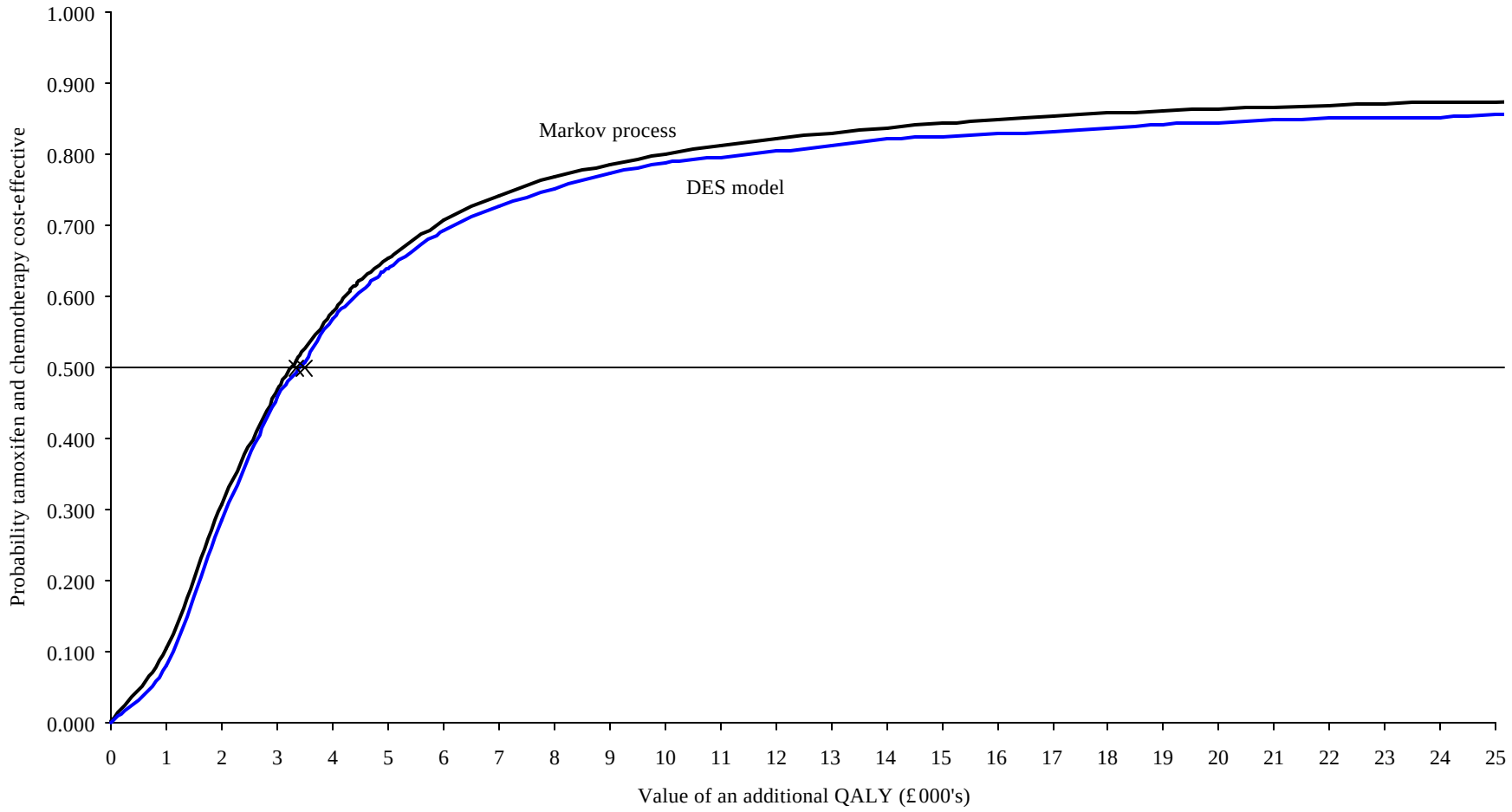


Figure 4 Cost-effectiveness acceptability curves derived using alternative modelling techniques



which the probability of tamoxifen and chemotherapy being the cost-effective therapy reaches 0.5 are £3,250 and £3,400 for the Markov process and DES model, respectively. The remaining portion of the CEAc curve shows that the probability of positive net benefits levels out at around 0.9 and 0.88.

Discussion

The alternative sets of results derived from the different modelling techniques show that the Markov process provides lower estimates of each model output, though ICERs are similar. The closeness of the results suggests that it is unlikely that the use of one model's results over the other would lead to an alternative resource allocation decision.

The lack of any significant difference in the results produced by the alternative models does not necessarily mean there are no significant differences between the two modelling techniques. During the course of developing and analysing the two models presented in this paper the author uncovered several potentially important differences between the two techniques.

In the following sections, four statements are made with respect to potential factors that could influence the choice of decision modelling technique. The four statements are discussed in the light of the work completed in this paper.

1. If model parameters are a function of the time spent in particular states DES will capture a more accurate reflection of the true relationships between health states.

The description of patient pathways was limited in the Markov process in one area of the model. Patients experiencing a locoregional relapse following DFI progressed to a more severe site of relapse (metastases) or straight to death. The data relating to the progression from locoregional relapse mainly presented disease-free survival (DFS) curves representing the probability of progressing in successive time periods.

To apply data from a survival curve to patients within a decision model it is necessary to record when separate patients enter the state of interest in order to apply differential

probabilities of experiencing an event over time. The Markovian assumption prevents the application of differential probabilities to patients within the same health state. In effect, a patient is a patient is a patient in every state within the model.

The Markovian assumption means that the data available in the survival curve must be transformed to a constant probability that is applied to every patient within the health state. The ideal output from the conversion of the survival data is a mean length of disease-free survival, which would produce, in the absence of discounting[2], the same mean results as those derived from the use of the survival curve directly. However, if the surviving proportion of patients presented by the survival curve does not reach zero, the conversion of survival curves to constant probabilities in the Markov process is approximate because mean 'survival' cannot be fully estimated.

In the DES model, patients remaining disease free at the end of the available follow-up were assumed to follow the survival profile of the general population. This is possible because not only does the DES model note the time at which patients enter the locoregional relapse health state, it also records the age of each patient. This enables the application of age-specific mortality rates to each patient that survives beyond the last point of follow-up reported in the survival curve.

In a Markov process there is no way of determining the age of the patient at the end of the period of follow-up. A single time period spent in the health state was required for the whole patient group. The distribution of DFS was significantly skewed so the median DFS was an inappropriate proxy for the mean, so some heroic assumptions were employed to estimate a mean period of DFS.

The representation of disease-free survival as a constant probability of experiencing an event also had another consequence. In the DES model it was possible to make the assumption that if patients had not experienced a relapse after 11 years in remission they were 'cured' and the next event of interest to the decision model was death. In the Markov process it was not possible to split the destination of patients according to

[2] The impact of discounting on the results derived from a DES model and Markov processes is discussed in the following section.

time spent in the state, so all patients leaving the state were subject to the same probabilities of experiencing a more severe relapse, or progressing straight to death.

To demonstrate the impact of this factor the Markov process was analysed employing the median DFS and the estimated mean DFS. The results are presented in table 3 showing that the effect on the aggregate costs is negligible, but there is a more substantial effect on the total QALYs and lifeyears associated with both therapies.

Table 3 Comparison of model outputs from the Markov process using the median length of disease-free survival and the approximate mean disease-free survival

	Costs	QALYs	Lifeyears
Mean (tamoxifen+chemotherapy)	£8,862	11.62	15.72
Median (tamoxifen+chemotherapy)	£8,934	11.23	15.13
Difference	-£72	0.39	0.58
Mean (tamoxifen alone)	£6,721	11.07	14.90
Median (tamoxifen alone)	£6,741	10.70	14.33
Difference	-£20	0.37	0.57
ICER (mean)	£3,896		
ICER (median)	£4,159		
Difference	£263		

At present both models employ the median survival times from each of the metastatic sites, which, as demonstrated above, is not the ideal data for such events. If survival curves become available to describe time to events in these states, the assumptions imposed on the data employed in Markov processes could considerably undermine confidence in the model's outputs.

2. If the specification of similar health states that differ only with respect to the experience of previous states compromises the clarity of the model, the use of DES should be considered.

The statement refers to the possible proliferation of health states in a Markov process in order to represent states that have similar characteristics, but the subsequent patient pathways are influenced by a patient's treatment history.

The flexibility of the DES approach was not used to its fullest potential in the ABC model due to data constraints. One possible area in which greater advantages over the Markov process could have been established is in the conditional development of

patient pathways. For example, DFI is known to be a prognostic factor that reflects the intrinsic growth rate of the tumour [Borner et al, 1994]. It may be plausible to assume that patients with short lengths of DFI will be more likely to experience a relapse, rather than dying with no evidence of disease, than patients with a longer period of DFI. Moreover, it has been reported that the length of DFI affects a patient's prognosis from the point of diagnosis with a recurrence [Ingle et al, 1994;Wong and Henderson, 1994].

The Markovian assumption would have a greater impact on the comparison of the two modelling techniques if more detailed data had been available for the evaluation. Unfortunately the secondary data sources used in this thesis did not produce data that described possible relationships in enough detail to warrant inclusion in the models. If primary data becomes available it is possible that certain patient characteristics could be employed as attributes within the DES model, which may provide an improved representation of the treatment area.

A related issue that did occur in the ABC models involved the description of the experience of toxicity. Three forms of toxicity were described - major, grade 3/4, and grade 1/2 - which could be experienced simultaneously. The available data only described the proportion of patients experiencing each form of toxicity. After consulting the clinicians involved with the study, it was decided that the most appropriate approach was to subject each patient to a probability of experiencing each form of toxicity.

In the DES model no separate health states were defined to represent toxicity, rather the experience of toxicity was assigned to each patient within the health state 'disease-free interval'. The attributes containing the information on the experience of toxicity were incorporated into the estimation of the costs and QALYs associated with each patient time in DFI as they left the state.

In the Markov process it would have been possible to create seven states that described each possible combination of toxicity states, though this would have required probability estimates for each combination of toxicities, which were not available in the literature. Hence, patients could only experience one form of toxicity

and if the sum of the probabilities exceeded 1, then the probability of experiencing the least severe form of toxicity (grade 1/2) was reduced.

The impact of the differential methods employed to represent the experience of toxicity are hard to quantify, though the author believes that the impact on the second-order simulation results were small because the mean experience of toxicity for groups of patients would only differ slightly in the cases where the probability of grade 1/2 toxicity needed adjusting so that the sum of the probabilities equalled 1.

3. If the format of the data required to accurately describe the timing of events within a model does not conform to that enabled by Markov processes then DES will provide a truer representation of reality.

Issues around the format of the available data are mainly applicable to modelling studies that employ secondary clinical data. In such circumstances the analyst is normally bound by the presentational norms for particular types of data. The process of modelling adjuvant therapies for breast cancer highlighted this issue in relation to the representation of survival from diagnosis with metastases.

The data describing survival were presented as set survival times, rather than data describing the probability of experiencing an event in a particular interval. The DES model permitted the author to incorporate such data in exactly the format that the data were available. Within each run of the DES model, for each sampled survival time, every patient experiencing a form of metastases remained in that health state for the sampled survival period.

The Markov process required the conversion of set survival times to constant probabilities of experiencing death at any point following the diagnosis of metastases. The following formula was employed to transform set survival times to constant probabilities:

$pr(event) = 1 - \exp\left(\frac{-1}{x}\right)$, where x is the set survival time.

To illustrate the impact on costs, QALYs and lifeyears, which are all subject to alternative discount rates, the author set up a macro in an excel spreadsheet to estimate the outputs associated with varying lengths of survival derived from the two models. The start year in the assumed state was also varied. Set survival times of between 1 and 40 months (accommodating the upper range of the survival estimates for soft tissue metastases), and start years of between 0 and 26 years were employed. To ascertain the impact on costs a discount rate of 6% was applied to monthly costs of £469 (the mean monthly cost of bone metastases). For QALYs, a utility value of 0.5 was discounted at a rate of 1.5%, whilst lifeyears were not discounted.

The ABC models covered a time horizon of 50 years, which is the maximum that the constant probabilities may be applied to patients remaining in a metastatic state. The macro linked the length of follow-up in the assumed metastatic state to the start year in the state. For example, if the start year in the state is sampled as year 10 then the outputs from the state are summed for the following 40 years in the state. This set-up provided the most accurate portrayal of the effect of converting set survival times to constant probabilities.

Figure 5a shows that the cost difference between the two modelling techniques, which is subject to the highest discount rate, is mainly affected by the length of survival, though the difference decreases as the start year increases. Figure 5b demonstrates that with a low discount rate the start year in a state does not affect the difference in QALYs greatly. For example, the difference in QALYs between the two models for a survival time of 40 months is 0.113 if patients enter the state in their first year in the model, a difference that decreases to 0.082 if the patients enter the state in their 26th year in the model. Figure 5c appears to show the most significant results, but the magnitude of the effects on lifeyears is small. The data illustrate that, in the absence of discounting, the conversion of set survival times to constant probabilities has little effect, though combinations of longer survival times and late entry into the assumed state start to produce larger differences.

Figure 1a Plot showing the difference in cost estimates between a DES model and a Markov process for survival from a state described as a set survival time, as a function of the survival time and the start year in the state.

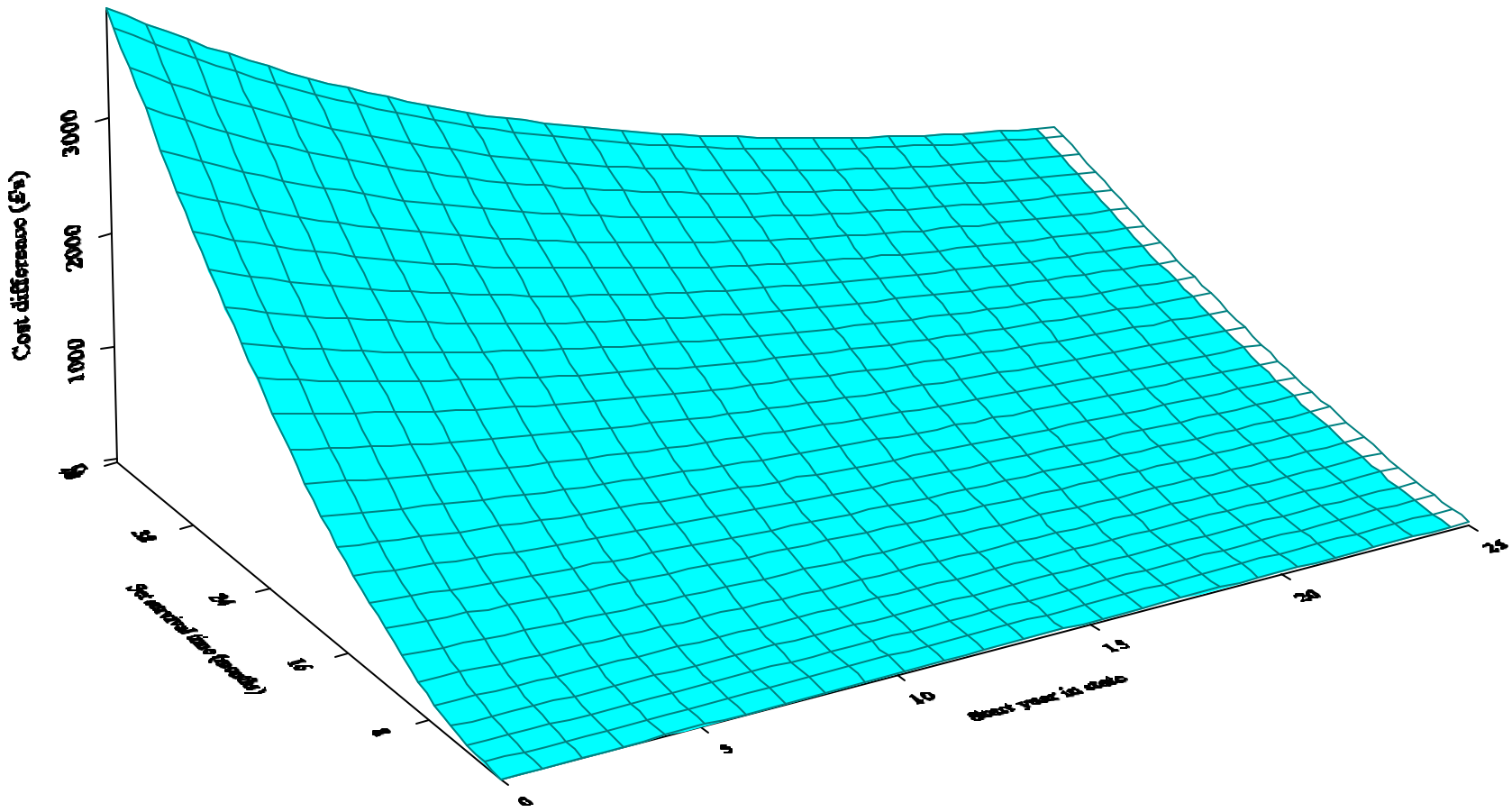


Figure 1b Plot showing the difference in QALY estimates between a DES model and a Markov process for survival from a state described as a set survival time, as a function of the survival time and the start year in the state.

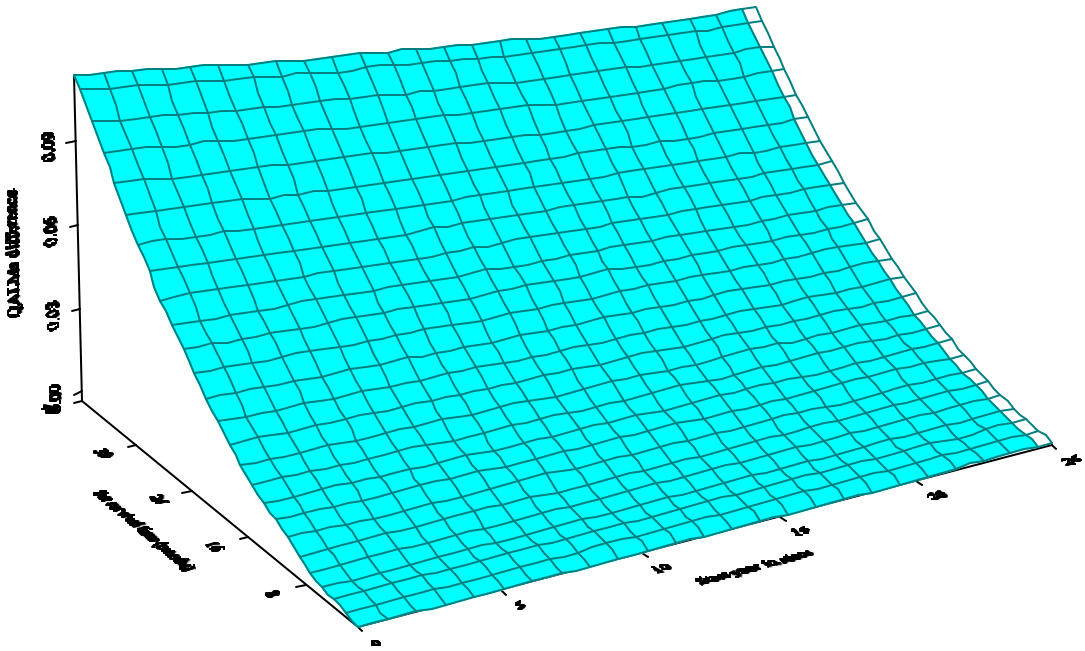
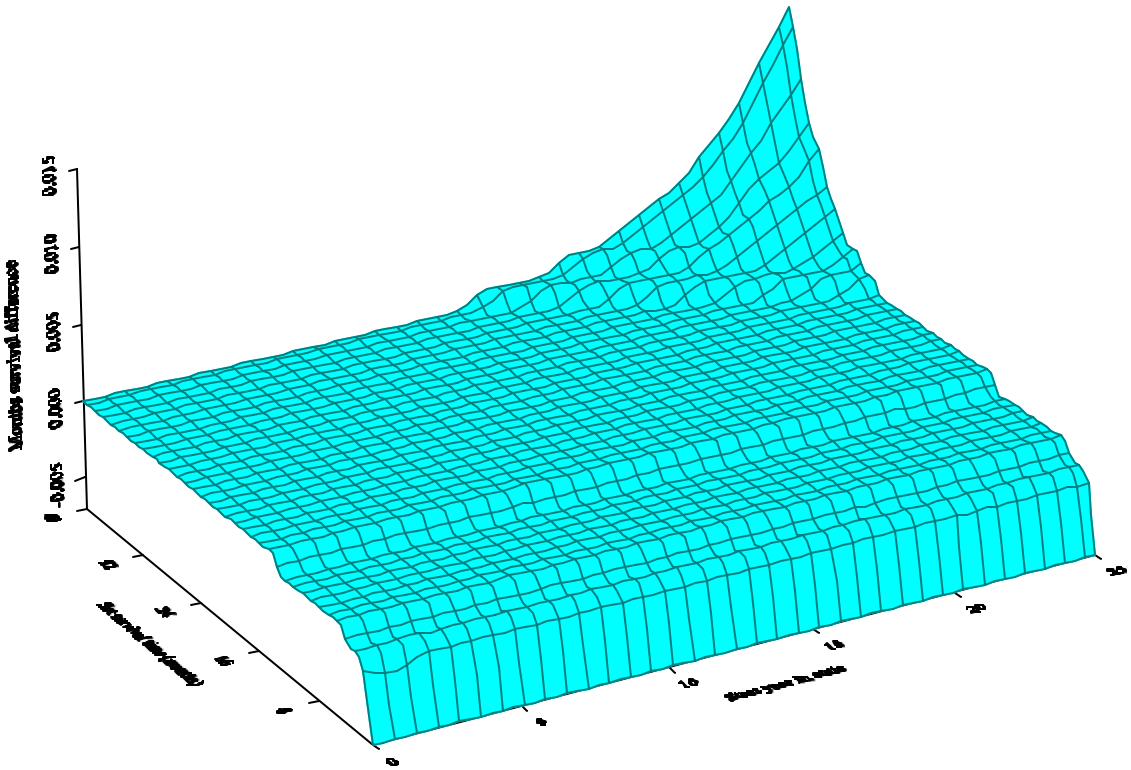


Figure 1c Plot showing the difference in lifeyear estimates between a DES model and a Markov process for survival from a state described as a set survival time, as a function of the survival time and the start year in the state.



The data show that the Markov process actually overestimates the relevant outputs at lower survival times, the maximum percentage overestimate is 8.2% for all three outputs (with a survival time of 1 month starting in year 0). Though the figures show that the effect of converting set survival times to constant probabilities diminishes as the start year in the assumed state increases, the extent to which the Markov process underestimates costs associated with higher survival times is extremely large in comparison to the overestimates. In addition, the threshold survival times at which the Markov process begins to underestimate costs and QALYs are low in the context of the survival times inputted into the ABC models.

Table 4 presents the maximum percentage over- and under-estimates associated with each of the model outputs for the Markov process, as well as the survival time at which the Markov process started to underestimate the outputs.

Table 4 Maximum over- and under-estimates of the ABC Markov process compared to the ABC DES model, and the threshold survival times for each model output

	Maximum overestimate (%)	Threshold for over-/under-estimates	Maximum underestimate (%)
Costs	£38 (8.2)	4 months	£3,812 (22.79)
QALYs	0.00342 (8.2)	5 months	0.113 (7.15)
Lifeyears	0.00683 (8.2)	26-39 months*	0.0155 months (0.46)

* the Markov process first overestimates lifeyears for a survival time of 26 months, but continues to underestimate lifeyears at high start years upto a survival time of 39 months. However, the magnitudes are very small, amounting to under 0.02%.

Between 50% and 60% of patients experience a metastatic relapse in the ABC models. The differential outputs caused by the conversion of set survival times to constant probabilities in the Markov process is likely to be a primary factor explaining differences in the aggregate outputs of the two modelling techniques.

4. If DES appears to offer a higher quality model, but the necessary inputs, in terms of expertise and time, are not available for the construction and analysis of a DES model then a Markov chain will have to suffice.

This statement is a truism; the lack of available analytic input necessarily precludes the use of the more complicated modelling technology. However, it is important to be able to estimate the required inputs prior to selection of the appropriate technique in order to make an informed choice.

The author is a health economist and this paper was based on the premise that the health economist wished to retain control of the economic HTA decision model, which meant that the modelling techniques adopted were accessible to analysts with no formal training in operations research. The author chose to analyse the Markov process as a cohort analysis because this was the most common form of analysis in the economic evaluation literature, but also because the more complicated first-order Monte Carlo simulation approach does not add to the analysis.

The DES model was built using software employed specifically to create DES models[Simul8, 2000]. The author accepted that DES models could be built using independent programming languages, but felt that an initial comparison of DES with Markov processes was best undertaken using software that provided guidance to the use of DES and facilitated the easiest construction of such models.

The overall time spent obtaining the final results from the DES model far exceeded the time employed using the Markov process. To present the causes of the excess input, the modelling process is split into two stages - building and analysing - each of which required differential elements of time and expertise between the two modelling techniques.

Building the model

A Markov process simply multiplies the proportion of patients in each health state in each time period by the associated cost and utility value attached to each state, applying the relevant discount rates to each time period. To estimate their average value, the outputs are then summed across all time periods.

In the absence of discounting and the use of attributes the building of a DES model would not be significantly more complex than a Markov process. To incorporate such activities within the model, the building of the DES model employs programming code, albeit relatively simple code. In terms of economic HTA decision models, the ABC model was a complex model because it covered a long time horizon and incorporated the use of patient attributes. The development of the ABC DES model

was a prolonged exercise, though during the course of building the model the author consistently discovered more efficient means of programming. Such advances are part of the learning curve that makes the modelling process shorter and less complicated.

Analysing the model

The most visible source of additional time required for the use of a DES model, rather than a Markov process, is during the analysis of the models. DES, being event-orientated, is necessarily analysed using first-order Monte Carlo simulations. This means that for each set of input parameter values, data on a large number of individual patients are collected to inform a mean value for each output. After experimentation, 10,000 patients were found to provide an adequate representation of the mean outputs for each first-order simulation of the ABC DES model, which was a function of the complexity of the model. Employing cohort analysis, the Markov process estimates the model's outputs for the whole cohort simultaneously.

The final analysis of the two models took around 1 hour and 3 days for the Markov process and DES model, respectively. During experimentation of the DES model it was necessary to ask the software to remember every random number that had been employed in previous runs of the model, in order to prevent the exact same sets of input parameter values being sampled repeatedly. This requirement slowed the analysis of the DES model considerably as the number of runs undertaken increased.

The author undertook 10,000 simulation of the Markov process, but only 2,500 runs of the DES model due to the cumbersome nature of the model. 2,500 runs provided adequate representations of the distributions of the model's output, though obviously 10,000 estimates provided better definitions of the distributions.

The time to analyse not only includes the final 'correct' experimentation with the models, but the whole process of verification and validation, which often requires significantly more analysis time than the final process of experimentation.

Space limitations prevented the presentation of the process of validation in this paper, but part of the data employed to validate the ABC models described economic outputs from the model for the first 4 and 10 years following diagnosis[Wolstenholme J and Whynes D, 1998;Castiglione-Gertsch et al, 1994;Cummings et al, 1993]. The process of validation using these data required collecting data for all patients at exactly 4 and 10 years, respectively, after they entered the model. This caused particular difficulty for the DES model.

To collect such data using the Markov process the model was simply stopped at the relevant period and the associated costs and effects collected. This is possible because all patients start at the same time in the model and the outputs are collected at regular time intervals – every month.

The collection of individual patient data for such specified time periods was more complicated in the DES model because it is event-orientated, rather than time orientated. The collection of data within the model is also event orientated. As a patient leaves a particular state the programming code looks back at the history of the patient within the state she is leaving and assigns the relevant costs and utility values.

The DES model could not simply cut-off the data collection period for patients at a specified length. In order to collect data constrained by time it was necessary to run the full model, but collect data at the end of each state up to the end of the specified time period. This was only possible by inserting programming code that effectively asked if the patient had been in the model for the specified period at regular intervals when aggregating the costs and effects at the point of exit of each state.

In effect, the process of validation was subject to a separate process of verification; to ensure that mechanisms put in place for the validation were working properly.

Conclusions

Despite the relative closeness of the aggregate cost-effectiveness results derived from the two ABC models, the structural differences between the models were potentially important. It would appear to be a matter of good (or bad, depending on your

perspective) fortune that the divergences between the models acted in opposite directions that almost cancelled each other out.

As the outputs estimated by the DES model were uniformly higher than those produced by the Markov process, it appears that the most influential sources of divergence between the ABC models were the alternative methods of inputting set survival times into the model. Next, the more flexible approach to describing survival curves outside the initial state in the model improved the validity of the DES model over the Markov process. Finally, the use of attributes in this particular treatment area probability had only a minor differential effect between the two models. A summary of the differential model characteristics is provided in the appendix.

On the basis of these results, the author proposes that conclusions can be drawn for the applications of decision modelling techniques to the evaluation of alternative treatment areas. The choice of decision model should be judged on the characteristics of alternative treatment areas, but if the necessary resources are available for the application of DES, it should be strongly considered if any of the three issues discussed above appear relevant. The relative importance of each of the three issues is likely to differ between treatment areas, but on the basis of the work presented in this paper analysts should be able to make an informed choice with respect to the appropriate decision modelling technique to employ.

Word count 7,000

Acknowledgements

I am grateful to the following colleagues for their generous assistance: Andy Briggs, Liz Fenwick, Karl Claxton, Jackie Brown, Martin Buxton and Mark Sculpher.

Reference List

Crystal Ball. 2000. Decisioneering, Inc. USA,

Simul8. 2000. Visual Thinking International Ltd. Glasgow, Scotland

Barton P, Robinson S, Bryan S (2000) The use of modelling in the economic evaluation of health care. *Health Economists' Study Group, Nottingham*

Borner M, Bacchi M, Goldhirsch A, Greiner R, Harder F, Castiglione M, Jungi WF, Thurlimann B, Cavalli F, Obrecht JP, et al (1994) First isolated locoregional recurrence following mastectomy for breast cancer: results of a phase iii multicenter study comparing systemic treatment with observation after excision and radiation. Swiss group for clinical cancer research. *Journal of Clinical Oncology*

Briggs A, Fenn P (1998) Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane. *Health Economics* 7: 723-40

Briggs AH (1999) A bayesian approach to stochastic cost-effectiveness analysis. *Health Economics* 8: 257-61

Briggs AH (2000) Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics* 17: 479-500

Buxton MJ, Drummond MF, Van Hout BA, Prince RL, Sheldon TA, Szucs T, Vray M (1997) Modelling in economic evaluation: an unavoidable fact of life. *Health Economics* 6: 217-27

Campbell H, Karnon J, Dowie R (2000) Cost analysis of hospital-at-home initiative using discrete event simulation. *Journal of Health Services Research and Policy* 5:

Cancer Research Campaign (1998) *Fact sheet 1: Incidence*. CRC Publications: London

Castiglione-Gertsch M, Johnsen C, Goldhirsch A, Gelber RD, Rudenstam CM, Collins J, Lindtner J, Hacking A, Cortes-Funes H, Forbes J, et al (1994) The international (ludwig) breast cancer study group trials i-iv: 15 years follow-up. *Annals of Oncology* 5: 717-24

Chaussalet T, El-Darzi E, Millard PH (1999) Markov and discrete event simulation models for the economic evaluation of alternative options for dementia services. *www.wmin.ac.uk/~chausst*

Cummings FJ, Gray R, Tormey DC, Davis TE, Volk H, Harris J, Falkson G, Bennett JM (1993) Adjuvant tamoxifen versus placebo in elderly women with node-positive breast cancer: long-term follow-up and causes of death [see comments]. *Journal of Clinical Oncology* 11: 29-35

Fairclough DL (1997) Summary measures and statistics for comparison of quality of life in a clinical trial of cancer therapy. *Statistics in Medicine* 16: 1197-209

Felli JC, Hazen GB (1998) Sensitivity analysis and the expected value of perfect information. *Medical Decision Making* 18: 95-109

Fox KR (1991) Adjuvant therapy of node-positive operable breast cancer. In *Breast cancer treatment. A comprehensive guide to management*, Fowble B, Goodman RL, Glick JH, Rosato EF (eds) Mosby Year Book: St Louis

Glick JH (1991) Adjuvant therapy for node-negative breast cancer. In *Breast cancer treatment. A comprehensive guide to management*, Fowble B, Goodman RL, Glick JH, Rosato EF (eds) Mosby Year Book: St Louis

Hillier FS and Lieberman GJ (1995) *Introduction to operations research.* McGraw-Hill: New York

Ingle JN, Foley JF, Mailliard JA, Krook JE, Hartmann LC, Jung SH, Veeder MH, Gesme DH Jr, Hatfield AK, Goldberg RM (1994) Randomized trial of cyclophosphamide, methotrexate, and 5-fluorouracil with or without estrogenic recruitment in women with metastatic breast cancer. *Cancer* 73: 2337-43

Iverson GR (1984) *Bayesian Statistical Inference.* Sage.

Lord J, Asante MA (1999) Estimating uncertainty ranges for costs by the bootstrap procedure combined with probabilistic sensitivity analysis. *Health Economics* 8: 323-33

Pasta DJ, Taylor JL, Henning JM (1999) Probabilistic sensitivity analysis

incorporating the bootstrap: an example comparing treatments for the eradication of helicobacter pylori. *Medical Decision Making* 19: 353-63

Pidd M, (1989) *Computer modelling for discrete simulation*. Wiley: Chichester

Rice JA (1995) *Mathematical Statistics and Data Analysis*. Duxberry Press: Belmont, California

Sonnenberg FA, Roberts MS, Tsevat J, Wong JB, Barry M, Kent DL (1994) Toward a peer review process for medical decision analysis models. *Medical Care* 32: JS52-64

Wakker P, Klaassen M (1995) Confidence intervals for cost-effectiveness ratios. *Health Economics* 4: 373-82

Wolstenholme J, Whynes D/ (1998) The costs of treating breast cancer: implications for screening. *International Journal of Technology Assessment in Health Care* 14:

Wong K, Henderson IC (1994) Management of metastatic breast cancer. [Review] [87 refs]. *World Journal of Surgery* 18: 98-111

Appendix

The respective assumptions incorporated in each model are presented in table A1.

Table A1 Respective assumptions employed in the Markov process and the DES model

Area	Markov process	DES model
General	Cycle length of 1 month.	Minimum time period of 1 month.
	Maximum age of patients 100 years old.	Maximum age of patients 100 years old.
Disease free interval	Annual probability of leaving DFI informed by survival curve.	Annual probability of leaving DFI informed by survival curve.
	Annual probabilities converted to monthly probabilities assuming constant transition rates.	Annual probabilities converted to monthly probabilities assuming constant transition rates.
Toxicity	Patients can experience one of three categories of toxicity.	Patients can experience any combination of three categories of toxicity.
	Grade 1/2, 3/4 and major toxicity end after 6 months.	Grade 1/2, 3/4 toxicity end after 6 months, length of major toxicity sampled from probability distribution.
Locoregional relapse	Patients remain in locoregional relapse for exactly 1 month.	Patients remain in locoregional relapse for exactly 1 month.
Remission	A mean time spent in remission employed.	Annual probability of leaving remission informed by survival curve.
	Patients subject to a constant monthly probability of leaving the remission state.	Annual probabilities converted to monthly probabilities assuming constant transition rates.
Metastases	A mean time spent in metastatic states employed.	A mean time spent in metastatic states employed.
	Patients subject to a constant monthly probability of dying.	Patients remain in metastatic states for set amount of time.

Table A2 presents a summary of the differences observed between the ABC DES model and the Markov process.

Table A2 Summary of impact of differences between the Markov process and DES ABC models

Model difference (in DES model)	Impact on DES results (relative to Markov process results)
Simultaneous experience of separate toxicity categories in the DES model	Increase in costs for patients receiving tamoxifen and chemotherapy because combined probability of toxicity can be greater than 1.
More flexible description of disease free survival after experience of a locoregional relapse	The mean length of DFS is longer in the DES model, but in the Markov process all patients leaving remission have a probability of experiencing a further relapse (metastases) because it is impossible to distinguish between patients within the health state 'remission'. In the DES model patients who remain disease free after 11 years are assumed to be subject to the mortality rate in the general population and go straight to death from remission. It is likely that this will lead to lower costs in the DES model, though the effect on lifeyears and QALYs is less clear.
Survival from metastases specified as set survival times	The Markov process underestimates all model outputs when survival times over 14 months are converted from set survival times to constant probabilities. The magnitude of the effect is influenced by discount rates. Assuming discounts rates of 6%, 1.5% and 0% for costs, QALYs and lifeyears, respectively, the largest impact is on the costs estimated by the model. The likely aggregate impact on the ABC models is that the DES model will produce higher estimates of all three outputs, but the effect will be greatest for costs, then QALYs and finally lifeyears.
Analytic input	There is a learning curve associated with the use of DES, though the use of specialist software reduces its' steepness. The analysis of DES models is substantially longer than for a Markov process. The length of analysis may be partly due to the use of specialist software.