



A taxonomy of model structures for economic evaluation of health technologies

Alan Brennan^{a,*}, Stephen E. Chick^b and Ruth Davies^c

^a*School of Health and Related Research (SchARR), University of Sheffield, Sheffield, UK*

^b*Health Management Institute, INSEAD, Technology and Operations Management Area, Fontainebleau, France*

^c*Operational Research and Systems, Warwick Business School, University of Warwick, Coventry, UK*

Summary

Models for the economic evaluation of health technologies provide valuable information to decision makers. The choice of model structure is rarely discussed in published studies and can affect the results produced. Many papers describe good modelling practice, but few describe how to choose from the many types of available models. This paper develops a new taxonomy of model structures. The horizontal axis of the taxonomy describes assumptions about the role of expected values, randomness, the heterogeneity of entities, and the degree of non-Markovian structure. Commonly used aggregate models, including decision trees and Markov models require large population numbers, homogeneous sub-groups and linear interactions. Individual models are more flexible, but may require replications with different random numbers to estimate expected values. The vertical axis of the taxonomy describes potential interactions between the individual actors, as well as how the interactions occur through time. Models using interactions, such as system dynamics, some Markov models, and discrete event simulation are fairly uncommon in the health economics but are necessary for modelling infectious diseases and systems with constrained resources. The paper provides guidance for choosing a model, based on key requirements, including output requirements, the population size, and system complexity. Copyright © 2006 John Wiley & Sons, Ltd.

Received 24 May 2005; Accepted 30 March 2006

Keywords health technology assessment; cost-effectiveness analysis; modelling methodology; simulation; decision tree; Markov model

Introduction

Decision modelling is used in many countries across the world to consider the cost-effectiveness of alternative healthcare interventions [1]. Agencies such as the UK National Institute for Health and Clinical Excellence (NICE) demand a systematic, evidence-based analysis of the costs and benefits of new technologies. This often requires synthesis of evidence in computer-based mathematical models. The structure of these models is

not pre-specified by decision makers. It is chosen by the analyst to best reflect the needs of the particular decision problem. The analyst must understand the healthcare system and the available evidence and must know which modelling structures are available and appropriate in which circumstances. If incorrect assumptions or structures are used then the model results will be flawed and the decision may be incorrect, with potential consequences measured not only in bad advice, or fluctuations in sales for a pharmaceutical company, but also in wasted resources in the

*Correspondence to: School of Health and Related Research (SchARR), University of Sheffield, Regent Ct., 30 Regent St., Sheffield, S14 DA, UK. E-mail: a.brennan@sheffield.ac.uk

healthcare system and ultimately in quality of life and survival in the population concerned.

The decision about which model structure should be chosen in a particular health economic evaluation context is only rarely discussed in published studies. Policy recommendations based on a model may depend upon the explicit and implicit assumptions of the model. Many evaluations use aggregate or 'cohort' models, which examine the proportions of the population undergoing different events with associated costs and utilities. Patient or individual level models, which sample individuals with specific attributes and follow their progress over time, have recently become more prevalent. The problem is that aggregate models adopt assumptions that may unwittingly produce inaccurate or inadequate solutions, whilst individual level models can adopt less stringent assumptions, but may be more time consuming to develop and run. Recent debates on choice of model structure have focussed on this cohort versus individual level dichotomy [2]. This paper presents a taxonomy of a wide range of available model structures, and guidance for the selection of an appropriate model.

We define a model-based evaluation as a formal quantified comparison of health technologies, synthesising sources of evidence on costs and benefits, in order to identify the best option for decision makers to adopt. Interventions can include drugs, surgical procedures, and psychological therapies or much broader health system strategies such as screening policies, public health interventions or bio-terrorism defence. Outputs usually include intervention costs and a measure of morbidity and mortality such as quality of adjusted life years (QALYs). Adoption may depend upon a single measure, e.g. the expected incremental cost per QALY between technologies against a threshold willingness to pay (λ) or equivalently, expected monetary net benefit ($NB = \lambda * QALY - Cost$). Sometimes, decision makers are also interested in other measures around uncertainty, individual or geographic variability and trends over time. Recent guidance recommending probabilistic sensitivity analysis [3, 4] leads to a preference for fast, efficient models with a single output measure.

Model structure is usually determined by considering the relationship between the inputs (natural history of disease, clinical pathways, evidence of interventions' effectiveness, utilities associated with health states, intervention and

other costs, etc.), and the output measures required by the decision maker. Practical considerations also include availability of data, the background and skill of the researcher and the type of software available. As the choice is usually made relatively early in projects, it is often hard for model developers to abandon an initial structure and start again.

There are several papers whose guidance on good modelling practice distils down to a series of principles (e.g. transparent structure, appropriate and systematic use of evidence) rather than more detailed guidance on the structure that is appropriate [5–8]. Sculpher *et al.* [9] are slightly more detailed, recommending that model structure be as simple as possible, consistent with the stated decision problem and a theory of disease and not defined by data availability or health service inputs alone. They do not address the issue of which specific modelling techniques to use, but implicitly assume, in making their recommendations, the use of a cohort model. Some researchers have directly compared alternative model structures. Karnon's economic evaluation of breast cancer showed that an individual level discrete event simulation and a cohort Markov model could be 'tuned' to produce the same results [10]. Koopman *et al.* tried to understand how the explicit and implicit assumptions of different types of epidemiological models may influence decisions that are based upon model output, as an alternative to tuning the output of one model to the output of another [11]. Barton *et al.* reviewing methodology in decision trees, cohort Markov models, and individual level models [12], suggest that the choice between these three approaches depends upon: whether pathways could be adequately represented by probability trees, whether a Markov model would require an excessive number of states, and whether interaction between patients is important.

This paper extends those articles by more systematically establishing the range of model structure options and criteria for their selection.

A taxonomy of model structures

Table 1 shows the range of available approaches for health economic evaluation models, and their relationships to each other. This section describes the taxonomy from a conceptual perspective.

Table 1. Taxonomy of model structures

| | | | A | B | C | D |
|---|------------------------|-----------------|--|---|---|--|
| | | | Cohort/aggregate level/counts | | Individual level | |
| | | | Expected value, continuous state, deterministic | Markovian, discrete state, stochastic | Markovian, discrete state, individuals | Non-Markovian, discrete-state, individuals |
| 1 | No interaction allowed | Untimed | Decision tree rollback | Simulated decision tree (SDT) | Individual sampling model (ISM): Simulated patient-level decision tree (SPLDT) | |
| 2 | | Timed | Markov model (evaluated deterministically) | Simulated Markov model (SMM) | Individual sampling model (ISM): Simulated patient-level Markov model (SPLMM) (variations as in quadrant below for patient level models with interaction) | |
| 3 | Interaction allowed | Discrete time | System dynamics (finite difference equations, FDE) | Discrete time Markov chain model (DTMC) | Discrete-time individual event history model (DT, IEH) | Discrete individual simulation (DT, DES) |
| 4 | | Continuous time | System dynamics (ordinary differential equations, ODE) | Continuous time Markov chain model (CTMC) | Continuous time individual event history model (CT, IEH) | Discrete event simulation (CT, DES) |

Appendix describes a variety of computational tools for implementing the models.

The rows (1–4) describe factors involving both time and interaction between individuals. Health economists' current approaches are largely those in the top half of the table. The models assume independence between individuals, and time may (row 2) or may not (row 1) be modelled explicitly. Models with interactions (rows 3, 4) are necessary when individuals interact (e.g. infectious disease transmission) or constraints affect individuals (e.g. finite service capacity or restricted supplies of organs for transplantation). We distinguish discrete time and continuous time models.

The columns (A–D) separate cohort from individual level models and disentangle assumptions concerning expected values, randomness, and the heterogeneity of entities. Cohort models (columns A, B) quantify the proportion of people with common characteristics. In cohort models with randomness (column B), the Markovian property is typically assumed, meaning that the future is conditionally independent of the past, given the present. Cohort models can account for different attributes/covariates (e.g. multiple ages,

weights, genders, other risk factors, stages of natural history of disease, etc.) by subdividing the number of states or branches, but the number of dimensions rises exponentially (e.g. M binary attributes imply 2^M dimensions). Individual level models (columns C, D) overcome this problem by simulating the progression of each individual with different characteristics (a population of N individual patients with M attributes each requires $M \times N$ data entries). Non-Markovian distributions (column D) allow greater flexibility in modelling the timing of health-related events. Non-Markovian distributions are those other than the exponential distribution (for continuous time models) and the geometric distribution (for discrete time models). Stochastic models may require many simulation runs to quantify the output mean and variance with sufficient accuracy. More complex individual level models (e.g. D4) can examine interactions both with other individuals and with the environment, including the availability of resources (e.g. doctors, beds, transplant organs). Each grid cell in the taxonomy is related to its neighbours by varying some of the basic assumptions that underlie each model.

Aggregate models without interaction

Decision trees for cohorts (A1, B1)

Decision trees (A1), as typically used in health economics, are among the most widely used aggregate level models [12, 13]. A decision tree outlines decisions (the square in Figure 1), the probability or fraction of various outcomes (emanating from the circles), and the valuation of each outcome, such as a QALY, cost or net benefit measure. The mean value of a decision is computed analytically ('rollback') by summing the probability of each outcome with its value, such as $E[\text{NB}|\text{Treat}] = 0.9 * 0.9 + 0.1 * 0.5 = 0.86$. This reflects either a large-population assumption, so that probabilities essentially match the actual fractions of individuals with a given attribute (a law of large numbers argument), or an assumption that expected values for patients are the desired outcome (rather than a distribution or variation, assuming a risk-neutral decision maker), or both. Recursion, or looping, is not allowed, but much more complicated scenarios are possible than in Figure 1, with many decision strategies, long sequences and multiple outcomes from chance nodes.

The simulated decision tree cohort (B1) provides an alternative approach to estimating the mean value of each decision option. In Figure 1, $E[\text{Number Cured} | \text{One million patient cohort treated}]$ can be estimated with Monte Carlo simulation by sampling a random variable with a Binomial (10^6 , 0.9) distribution, where 0.9 is the probability of a cure. By Monte Carlo, we mean the use of statistical sampling to approximate uncertainty in quantitative problems. This simu-

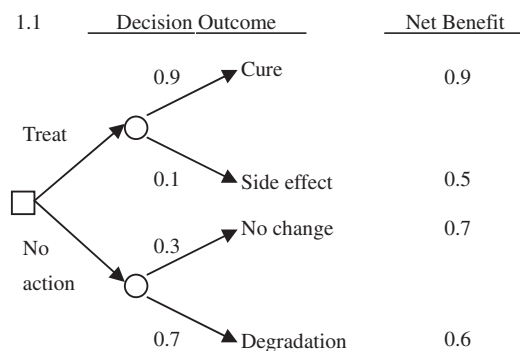


Figure 1. A decision tree

lates the number of individuals on each path of a decision tree, but not each individual separately. The advantage of this approach is that it can provide a measure of the variability of the number of individuals likely to be in each state.

Markov models for cohorts (A2, B2)

Markov models (A2) can provide a more compact representation than the decision tree when a repeated set of outcomes is possible through time. A cohort-based Markov transition matrix uses a transition probability per unit time for individuals in the cohort to change to another state, with associated costs and utilities [12, 14, 15]. The rows of the transition matrix must add to one (a conservation of probability mass). Figure 2 provides a sample transition probability matrix, as well as a state transition diagram that concisely describes the potential state changes graphically.

The usual formulation assumes that the transition probabilities are constant over time. Survival in a particular state is therefore defined by a geometric distribution in discrete time. Analysts often use subtle ways to extend the generalisability of the Markovian assumption, for example, by using different transition matrices as time progresses. If transition probabilities depend on certain attributes, then one can redefine states. For example, patient history can influence risk by defining states to include healthy with previous history of illness, healthy with no previous history of illness, etc.

It is a common misconception that analysts should use a time step defined by data availability (e.g. a year). The time-step used, however, affects the model results. With discrete time models, electing a smaller and smaller time slice is warranted until the output is no longer affected. To determine the probabilities for different time intervals for a two-state transition function, one can use an equation that relates the probability p_{ij} of an outcome, j , from state i through time t via the instantaneous hazard rate r_{ij} , namely $p_{ij} = 1 - e^{-r_{ij}t}$ [14]. Alternatively, the transition probability can be modelled by approximating p_{ij} by $r_{ij} * t$ for sufficiently small t (both have the same first-order linear term in their Taylor series expansion, as justified by references given in the discussion of discrete-time Markov chains below).

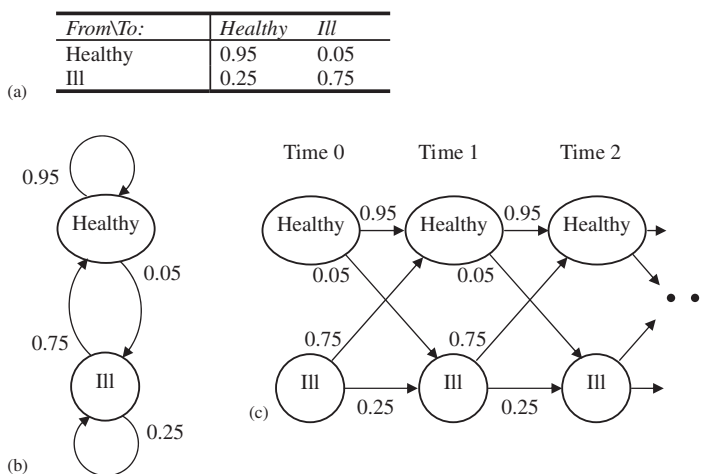


Figure 2. Several ways to specify a Markov model: (a) transition probability matrix for a Markov model; (b) state transition diagram; and (c) repetitive decision tree structure

Markov models for cohorts can be processed either analytically (A2) with expected values or with simulated random Markov model transitions (B2). The approach and rationale of Monte Carlo simulation is like that for the Simulated Decision Tree (B1).

```

For  $i=1, 2, \dots, n$  individuals
  Sample the patient attributes (e.g., risk factors)  $\theta$ 
  Determine the path of patient  $i$  through the model whose
  probabilities may depend upon  $\theta$ 
  Determine cost  $c_i$  and utility (QALY)  $u_i$  for individual  $i$ 
End For
Estimate the mean cost and utility  $E[(C U)]$  by  $\hat{\alpha}_n = \frac{1}{n} \sum_{i=1}^n (c_i, u_i)$ 
    
```

Figure 3. An individual sampling model algorithm

Individual sampling models without interaction (CD1, CD2)

Rather than tracking data for every compartment or path, individual sampling models (ISMs) track specific individuals. The individuals have potentially heterogeneous characteristics that affect their progression through the model, yet individuals progress through the model independently of each other and of environmental constraints. Typically, ISMs generate a large numbers of simulated patient histories and evaluate results with a sampling algorithm (Figure 3).

Simulated Patient Level Decision Tree models (CD1) can be used to simulate individuals through a tree's pathways, maintaining a record of the 'history' as nodes are processed. Time elapses implicitly. This approach is most useful when probabilities, QALYs, and costs depend upon a large or conceptually infinite number of values (e.g. normally distributed costs that depend upon patient glucose levels).

The most common implementations of ISMs in health economics are Markov models that are modified to simulate individuals through each time period (CD2). One key advantage lies in modelling multiple co-morbidities which depend on multiple attributes/covariates. Examples include models of diabetes where patient co-morbidities interact and affect the outcome [16–18], and other work on rheumatoid arthritis [19] and osteoporosis [20]. The ISM approach can be further modified to simulate the 'time to next event' rather using equal time periods. This can provide richer modelling power in applications where the timing of the response may influence the costs and benefits of the intervention (e.g. thrombolytic therapy for heart attacks, where an explicit model of the time of the response is critical to assessing the health benefit), computational efficiency and greater elegance in coding. For example, the Birmingham Rheumatoid

Arthritis Model [21, 22] considers long-term treatment for rheumatoid arthritis with strategies defined by alternative sequences of disease-modifying anti-rheumatic drugs.

Aggregate models with interaction allowed

System dynamic models (A3, A4)

If individuals do interact, it is not always necessary to model each individual separately. System dynamics, a common cohort-based approach in epidemiology and operational research, models the state of the system in terms of changing, continuous variables over time. Crucially, it enables the rate of change in the system to be a function of the system's state itself (i.e. feedback). Typical examples of feedback include infectious disease outcomes, where higher levels of infection produce higher risks of further infection but also reduce the number of people in the susceptible pool, and health care service constraints, where the system performs differently when it is full or over-capacity [23–25]. Factors affecting rates of change may include behavioural influences.

Continuous time system dynamics models (A4) use ordinary differential equations (ODE), like $dx/dt = f(x, t)$ to describe the rate of change of a state vector x . For example, Equation (1) describes the rate of change in the number of infected individuals i (the 'state') in a population of N individuals as a function of the contact rate (c contacts per unit time), the probability β of infection per potentially infectious contact, the probability i/N that one of the $(N-i)$ susceptibles contacts an infected, and the mean duration of infection D . This infection systems dynamics model is expressed as an ordinary differential equation as in Equation (1) below:

$$\begin{aligned} \frac{di}{dt} &= \text{population infection rate} \\ &\quad - \text{population recovery rate} \\ &= c\beta \frac{i}{N}(N-i) - \frac{i}{D} \end{aligned} \quad (1)$$

System dynamics models can have multiple linked ODEs to specify the full model.

The Euler-forward approximation to the continuous time ODE model results in a discrete time finite difference equation (FDE) model (A3) approximation

$$x_{j+1} = x_j + f(x_j, t_j)\Delta t$$

where x_j is the state of the system at time $t_j = j\Delta t$ [26]. The facilities for interaction are the same as for the continuous model. The solution to the differential or difference equations is specified over time, so the benefits that result from different health technologies can be quantified through time. The approach therefore provides the ability to model the deterministic interaction of groups of individuals but without modelling each individual in the system.

System dynamics models are commonly implemented in packages with graphical interfaces, which make them easier to build and use. There have been applications concerned with community care, short-term psychiatric care, HIV/ AIDS [24], smallpox preparedness [27–29], CJD [30] and others.

The cohort level Markov models (A2) that are evaluated using expected values are in fact special cases of discrete time finite difference (FDE) models (A3), but without the ability to model interactions (e.g. nonlinear dynamics from interpersonal infection transmission or resource capacity constraints). More accurate FDE approximations to ODEs are available [31], and partial differential equations [25] can complement ODEs to enhance model richness (e.g. geographic effects).

Discrete state, continuous time Markov models (B3, B4)

The system dynamics approach has two important assumptions – firstly, that the change dynamics are deterministic and secondly, that fractions of individuals can occur in specific health states (an infinitely divisible population assumption). With large populations such assumptions may be reasonable. Continuous time Markov chain (CTMC, grid cell B4) can address both a need to model an integer number of individuals in each health state and the variability that can occur in such systems. A CTMC can model many interactions including, for example, infectious disease dynamics and limited healthcare resources. Monte

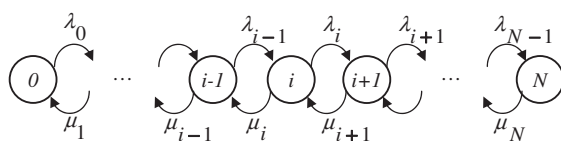


Figure 4. Continuous-time population Markov model transition probability

Carlo simulation is often employed to analyse these systems, although analytical stochastic process methods may sometimes be employed for sufficiently simple models.

A CTMC that is analogous to the system dynamics model in Equation (1) is depicted in Figure 4. Each circle is a potential state of the system (i people infected in a population of N individuals), the arrows out of a particular state represent the possible state transitions (from i people infected, a recovery changes the state to $i-1$ people infected, an infection changes the state to $i+1$ infected). The rates may depend upon the current state, so that the rate of a new infection occurring when there are i people infected is $\lambda_i = \beta i(N-i)/N$, and the recovery rate is $\mu_i = i/D$.

The distribution of the time to the next event (infection or recovery) is exponentially distributed with mean that is one divided by the sum of the rates out of the current state ($1/(\lambda_i + \mu_i)$). The probability of a particular event is proportional to the appropriate transition rate (an infection event has probability $\lambda_i/(\lambda_i + \mu_i)$). A CTMC in general can be simulated by sampling an exponentially distributed time to next event, then sampling the specific event type using probabilities that are proportional to the individual event rates. After sampling the specific event type and updating any statistics, the rates may change for sampling the time to the next event. The basic theory to support this sampling mechanism is Theorem 1.

Theorem 1

(Sampling time to first of many events). If T_1, T_2, \dots, T_k are independent random variables for the times to events 1, 2, ..., k with exponential distribution $T_j \sim \text{exponential}(\text{rate } \gamma_j)$, and mean $E[T_j] = 1/\gamma_j$, for $j = 1, 2, \dots, k$, then the time to first event $T_{\min} = \min\{T_1, T_2, \dots, T_k\}$ has exponential (rate $\sum_{m=1}^k \gamma_m$) distribution, and the probability that $T_j = T_{\min}$ is $\gamma_j / \sum_{m=1}^k \gamma_m$.

Proof: See, e.g. Çinlar [32]. Also see Banks *et al.* p. 222 [33].

In a context where k different events could happen (e.g. recoveries or potential infections of

many classes of ill people), Theorem 1 says that the time to the next event can be sampled first (exponential, rate $\sum_{m=1}^k \gamma_m$), and then the type of event to occur can be sampled next (event j with probability $\gamma_j / \sum_{m=1}^k \gamma_m$).

CTMC models are often analysed stochastically by simulating realisations, or sample paths, that describe how the state changes through time. The number of individuals in each state is tracked, and each individual event must be processed, but there is no need to sample the identities of individual patients to maintain their histories. Multiple independent realisations can be used to estimate statistics like time-varying means, time averages, and the probability of extreme events like outbreaks. If the number of individuals in the population gets very large, and a scaling is done so that the fraction of individuals in each state is maintained, then plots of sample paths get less variable and the number infected converge to a differential equation (e.g. see [34, 35] for epidemic models, and [36] for related results for queueing models). Figure 5 illustrates this by comparing the ODE infection model in Equation (1) with the stochastic analogue in Figure 4. The approximation is poorer for small populations in general. The upper left of Figure 5, where the possibility of infection being eliminated in a short time is high, gives a specific example of a poor approximation.

The discrete time Markov chain (DTMC) model (B3) is like the CTMC except that it is updated with finite time steps, at times $0, \Delta t, 2\Delta t, 3\Delta t, \dots$ for some suitably chosen period Δt . Brookhart *et al.* [37] used a DTMC to model a Cryptosporidiosis outbreak due to water treatment failure. The larger the period Δt of the DTMC, the worse a DTMC approximation of the original CTMC process becomes, in part because individuals are allowed to have only one event affect them per time step. The error in a DTMC approximation to a CTMC can practically be controlled by shrinking the time step size and rescaling parameters appropriately as described above for simulated Markov models (B2).

Individual level models with interactions

To model interactions, analysts often intuitively think of modelling individuals as separate entities.

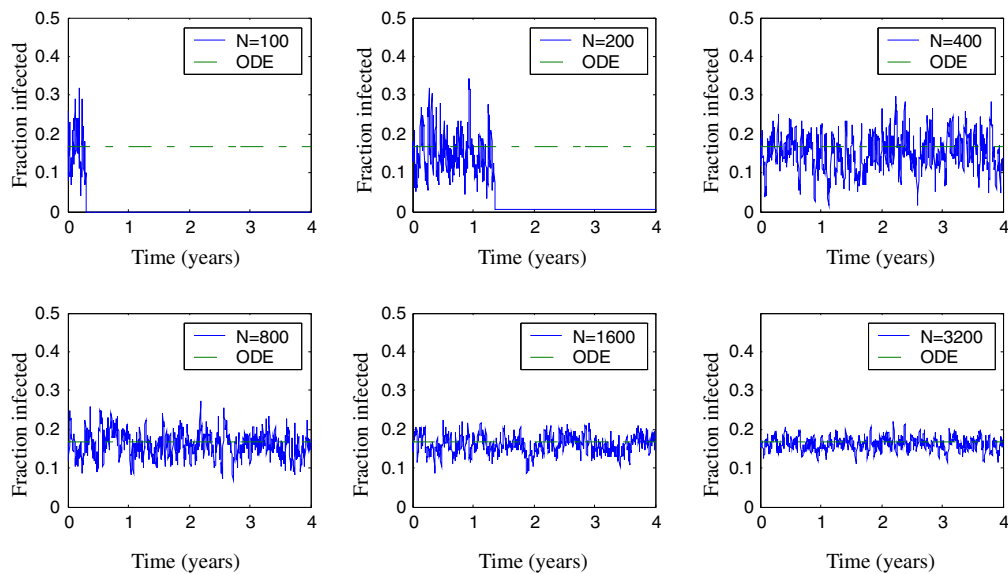


Figure 5. Scaled realizations of ODE infection model Equation (1) and stochastic model in Figure 4 (increasing population size N ; $c = 10$; $\beta = 0.1$; $D = 1.2$ days)

Individual level Markovian models with interactions (C3, C4)

Individual level Markovian models with interactions can be thought of as an extension of Markovian cohort models with interaction (B3, B4). Individual-level means that patient characteristics may be heterogeneous, and that histories may be tracked for each individual in the population. For that reason, this falls into the class of models that Koopman *et al.* calls individual event history (IEH) models.

There are at least two methods to simulate CT IEH models (C4). The first explicitly uses the Markovian assumption in Theorem 1 to simulate the time to the next event, followed by the selection of which event occurs. This differs from the use of the Theorem for Markovian cohort models (B4) in that a rate γ_{jl} is assigned for any event j that can occur for each individual l (rather than for each cohort). For the epidemic model in Figure 4, if event 1 is identified with infection, then the infection rate of susceptible individual l when there are i infected individuals is $\gamma_{1l} = c\beta(N - i)/N$. More generally, the parameters and rates may differ for each individual to reflect heterogeneous population characteristics, and

rates may also depend upon resource constraints (e.g. the service completion rate is nonzero for at most one patient if there is only one server). A second method is to use discrete-event simulation (see below) with exponential distributions.

The analogous discrete time IEH model (C3) steps forward in discrete time intervals in the same way as the DTMC model. Multiple individuals may be sampled at each time step, but each individual experiences at most one state transition per time step.

Discrete event simulation (D4, D3)

Probably the most flexible of all modelling techniques, continuous time discrete event simulation (CT, DES) (D4) describes the progress of individuals (entities), which undergo various processes (events) that affect their characteristics and outcomes (attributes) over time. Law and Kelton [38] present an excellent introduction. DES permits sampling from any distribution, Markovian or not. A queue structure enables interaction to take place with constraints and between entities. The state of the modelled system includes the current entities, their attributes, and a list of events that can occur either at the current simulation time

or that are scheduled to occur in the future. Events change the state at discrete times, and each event is processed one by one, in a non-decreasing time order. As each event is processed, it may produce consequent additional events to be processed at the current simulated time, or at some sampled time in the future.

The nomenclature DES has been applied differently in different papers in health economics [12], epidemiology, [39–41] and health care delivery system design or scheduling [42]. We define DES to be consistent with the operational research literature that developed DES theory and applications for the last 50 years [38, 43–45], and for which rigorous mathematical formalisms exist [46, 47]. In a healthcare DES, individual interacting patients are usually the entities. Additional entities may include health care service system resources, such as doctors, nurses, and ambulances for transport. Examples of disease models include models of the activities and non-Markovian distributions related to the natural history and treatment of coronary heart disease [48], of vertical transmission of HIV [49], and of vaccine trial design [39]. While constrained resources pose no problems for most DES tools, special care may be required to model infection dynamics or multiple correlated health risks. For example, models of multiple diseases may require multiple events on the future event list per patient, as well as a need to reschedule events if there are related underlying causes. POST is a tool to facilitate such event management [50]. Custom-built programs are sometimes produced to adapt to the special needs of DES models of infection [39, 51, 52].

Simulation models have sampling error for estimating mean output values (costs, QALYs, prevalence, etc.) that running multiple simulations can reduce. DES has the additional disadvantage of potential computational slowness associated with inserting and removing events from the future event list.

DES models have discrete-time analogues (D3) that are not different in any significant methodological way to the continuous time models if the discrete time steps are small. In fact, most commercial DES packages have animation graphics that show the entities moving through the system to aid model development and checking, and these typically implement a discrete-time DES variation if and when graphics are running.

How to decide which model structure to use

Several authors suggest selecting the simplest model that addresses the objectives of the study and structure of the disease and treatment process [3–11]. We agree with that recommendation, and propose the use of the taxonomy, which summarises the key structural assumptions that underlie each type of model, as a tool to help realise that recommendation.

Barton *et al.* provide a flowchart that determines which of five types of models is most appropriate, given the answers to four questions. Their first question is whether interactions are important. That determines whether one is in the first two rows or the last two rows of our taxonomy. Their second question, which applies when interactions are important, addresses whether cohort information (lower left quadrant of our taxonomy) or individual level information (lower right quadrant) is available. Our taxonomy refines and extends the set of available choices in that case. For example, Barton *et al.* mention system dynamics models to which we have added stochastic Markov chains. If interactions are not important, Barton *et al.* suggest assessing whether patient pathways are well-represented by decision trees, and if not, whether Markov models or individual sampling models are more appropriate. The taxonomy in Table 1 changes the question from whether decision trees are appropriate or not to the question of whether timing is important for cost and health outcomes or not, and whether mean values, randomness, and patient heterogeneity are important or not. The taxonomy suggests emphasising which features and assumptions of the decision problem are most important, as a means to enriching the set of potential models in Barton *et al.*

Table 2 gives a numbered checklist of additional key issues [I.1–I.15], which can help to identify an appropriate model from that enriched set of potential model types. The checklist follows the taxonomy structure, moving through the assumptions from the top left (the aggregate decision tree with homogeneous states, no recycling, no interaction, and a large enough population to make an integer approximation valid) across to the right, and down to the circumstances requiring a DES approach. When several model types are potentially appropriate, factors such as the available

Table 2. Issues and guidance on choice of model structure

| | Issue | Example | Choice of model |
|------|--|--|---|
| I.1 | Does the decision maker require knowledge of variability to inform the decision? | Effects of intervention are small and variable over time | Need for stochastic output (columns B–D) |
| I.2 | Is the decision maker uncertain about which sub-groups are relevant and likely to change his/her mind? | Decision maker may want to sub-divide the risk groups or test new interventions | Individual level models are more flexible to further covariates or changed assumptions (columns C–D) |
| I.3 | Is Probabilistic Sensitivity Analysis (PSA) required? | Decision maker uses cost-effectiveness acceptability curves or expected value of information | Deterministic model may be preferred (column A) but need for PSA should not drive model structure decisions |
| I.4 | Do individual risk factors affect outcome in a non-linear fashion? | Effects of age, history of disease, co-morbidity | Need to subdivide states in an aggregate model. Need to consider individual level modelling if the number is large. (columns C–D) |
| I.5 | Do covariates have multiple effects, which cause interaction? | Co-morbidities in diabetes affect renal failure and retinopathy | Individual level modelling likely to be necessary. (columns C–D) |
| I.6 | Are times in states non-Markovian? | Poor survival after an operation, moving from one age group to another, length of stay in hospital | Need to use ‘fixes’ in Markovian models or use non-Markovian models (columns D) |
| I.7 | Is the dimensionality too great for a cohort approach? | Large number of risk factors and /or subdivision of states to get over non-Markovian effects | Individual level modelling likely to be necessary. (columns C–D) |
| I.8 | Do states ‘recycle’? | Recurrence of same illness (e.g. heart attack, stop responding to drugs) | Decision tree approach is probably not appropriate (rows 2 to 4) |
| I.9 | Is phasing or timing of events decisions important? | In smokers, if lung cancer occurs before bronchitis, then patient may die before bronchitis occurs | Possible to have different branches in the decision tree but Markov model or simulation may be necessary. (rows 2 to 4) |
| I.10 | Is there interaction directly between patients? | Infectious disease models | Models with interaction (rows 3, 4) |
| I.11 | Is there interaction due to constrained resources? | Models with resource constraints | Models with interaction (rows 3, 4) |
| I.12 | Could many events occur in one time unit? | Disaster, outbreak of infection, risk of co-morbidities (e.g. diabetes) | Need for small time intervals or continuous time models (row 4) |
| I.13 | Are interactions occurring in small populations? | Use in hospital catchments area rather than nationally | Need to consider individual level modelling because of the inaccuracies in using fractions of individuals (columns C, D, rows 3, 4) |

Table 2. *Continued*

| | Issue | Example | Choice of model |
|------|--|---|--|
| I.14 | Are there delays in response due to resource constraints which affect cost or health outcome | Rapid treatment with angioplasty and stents after a myocardial infarction | Need for stochastic output and interaction (columns C, D, rows 3, 4) |
| I.15 | Is there non-linearity in system performance when inherent variability occurs? | A marginal change in parameters produces a non-linear change in the system ITU is suddenly full and newly arriving patients must transfer elsewhere | DES useful |

modelling software, implementation skills, and project time constraints must be balanced with the assumptions that underlie the chosen model. The key factors in selecting a specific model type delimit the main quadrants of the taxonomy.

Decision makers' requirements. The purpose of health economic evaluations is to inform decision-making. If the decision maker's perspective requires knowledge of the variability in the results rather than just mean outputs [I.1] then stochastic rather than deterministic models may be required. Such requirements might include knowledge of the range of impact achievable in different geographical or organisational contexts or at different times. For example, Davies *et al.* [53] show that, in screening for and eradicating *Helicobacter pylori*, the number of lives saved from death from gastric cancer may vary widely, meaning that some Health Authorities may incur large costs with little benefit [53]. It can be argued whether such variability should be part of the decision makers' considerations, but if it is, then the appropriate model structure is required.

Decision makers such as NICE now expect probabilistic sensitivity analysis (PSA) to quantify the uncertainty in mean outputs due to parameter uncertainty [I.3]. A deterministic model will provide the mean output for a given parameter set directly. In individual sampling models, the mean output for a given parameter set is estimated from (sometimes hundreds or thousands of) sampled individuals and the PSA will thus be relatively time consuming. The need for PSA should not directly drive model structure decisions but will influence them. There are interesting

developments to facilitate PSA including Gaussian meta-models of individual level models [54, 55] and quantifying the efficient number of individuals needed in each sample [56].

The choice of states and risk factors and the identification of their relationship to each other should normally precede the choice of model structure. There are occasions, however, when the decision makers require a very quick answer to a problem and there is a temptation to provide a model with a simple cohort structure without testing the underlying assumptions. At the other extreme, some decision makers require a model that can be adapted over time to examine different populations and include emerging evidence on new risk factors or treatments. In these circumstances, the analyst should consider the use of an individual sampling model and, in particular, discrete event simulation which is more flexible [I.4].

Sub-division into states. In all models the states represent the natural history of disease, treatments and their effects. In cohort models the states are often subdivided according to risk factors and may be further subdivided into short time periods to approximate non-Markovian distributions [I.6]. The representation of risk factors in a cohort model leads to the danger of an explosion in the number of states, discussed previously [I.7]. It is clearly desirable to keep the number of states as small as possible [9]. Risk factors may be combined if they have little impact on the results of interest to the decision maker.

The states in a cohort model are assumed to represent a homogenous population and the implicit use of averages by deterministic models may cause statistical bias in estimates of the mean

outputs. For example, if the heart attack risk (per unit time) increases with age, then the average of the heart attack rates in a population is higher than the heart attack rate at the average age. The larger the age groups, the worse this problem will be. This is a consequence of Jensen's inequality, which states that $E[f(X)] \geq f(E[X])$ when $f(\cdot)$ is a convex function (heart attack risk) and X is a random variable (a covariate like age) [32, 57]. More generally, if risk factors are not linearly related to the covariates, there may be some bias in QALY and cost measures. Individual level modelling can cope with a non-linear interaction between the risk and outcome [I.5], for example the relationship between the level of cholesterol and the risk of coronary heart disease [58].

Resource constraints. Another example of the inappropriate use of expected values is in the presence of variability and resource constraints [I.14]. Deterministic models that use average service times and mean demand rates can underestimate or completely miss queuing delay effects. If the event times are non-Markovian, then the queuing effects may differ even more (Pollacek–Khintchine formula) [57]. If health outcomes or delivery costs are linked to delays before intervention, stochastic models are required. Where there are explicit resource constraints and non-Markovian time to event distributions [I.15], discrete event simulation may be the best choice. Non-Markovian distributions are recommended as the most natural way to model process improvements in health care delivery, such as reductions in the variability in patient arrival and service delivery times [59].

Model run time and population size. Aggregate models sometimes need thousands or millions of states, possibly to cope both with the number of co-variates ('the curse of dimensionality') and to 'fix' non-Markovian effects. The resulting models may lack transparency and be slow to run. Individual sampling models can solve these problems, but on the other hand, need time consuming multiple replications to get good estimates of means. Discrete event simulation may be particularly slow because of the need to maintain a future events list. Variance reduction methods are available to reduce the number of replications, and hence the time needed [38].

The time taken by individual sampling models will increase with the size of the population to be

modelled whereas deterministic cohort models presume a law of large numbers effect. For large populations therefore, providing the non-linearity assumptions are met, a cohort model may be an appropriate choice.

It may be important, however, that the population remain large in individual states. For example, in infectious disease models, the stochastic effects of local die-out of infection in states with small numbers of individuals may result in very different dynamics from deterministic models [60]. For diseases that are transmitted primarily through smaller social structures (e.g. local families and local neighbourhood transmission), then stochastic models may be preferred since the implicit large population assumptions for deterministic models do not apply. Further, they show that stochastic models and ODE models may give very different outputs, or imply very different decisions and/or parameter estimates, when the basic reproduction number is low and when infection is transmitted in smaller social structures.

Generic models. The sponsor of the modelling study may specify the development of a generic model so that it may be adapted easily for new interventions or diverse populations. Sculpher *et al.* [61] suggest that multi-level regression models may help to understand the characteristics of different locations and how they may affect the model results. Cooper *et al.* [48] developed a model for the Department of Health, who required that the model assess both current secondary prevention drugs and future, as yet, unknown drugs. Here, the need for a comprehensive model which can be adapted to different requirements may override the desire for simplicity and speed of model development. However, developing or using an existing generic model needs careful consideration. For example, many influenza models exist for specific purposes: antigenic shift of the virus itself; viral dynamics in a human host; epidemic transmission from human to human in communities; or the effectiveness of global vaccination programs for pandemic preparedness. The development of a generic influenza model to handle all of these purposes would be untenable with existing or even envisioned future modelling technology. On the other hand, if a generic model is available for the decision, then it should be used, provided that model structure and the assumptions made are checked as appropriate in line with our taxonomy.

Discussion

This new taxonomy grid extends the palette of modelling approaches beyond those commonly used in health economics. This paper describes the underlying theory linking each approach and criteria for selection, and advocates the use of simple models that still accurately reflect disease progression and health care delivery to the extent needed by a given decision problem.

It is the responsibility of model developers, rather than policy makers, to select the most appropriate modelling approach. It is particularly important to think creatively about the techniques to be used where there are interactions and large uncertainties. This is because different modelling approaches can produce very different results; one example is the comparison of the different smallpox models that differ in how they model capacity constraints for vaccination and contact tracing, and random effects [27–29, 51]. Another application is water treatment policy to reduce the risk of human infection due to microbes in drinking water. Standard risk assessments assume independent risks across individuals that drink tap water (no interaction), but [62, 63] shows that a more complete model of secondary transmission among humans that are also associated with microbes in drinking water can result in very different policy recommendations regarding local interventions (special filters on the taps of immuno-compromised individuals) versus global interventions (municipal level treatment options).

The taxonomy grid is incomplete in the sense that further combinations of approaches and assumptions are also practised. A DES tool can track counts of individuals in various states rather than maintain data on each individual. The model of microbial drinking water infection combined ODE/CTMC models (for microbe concentration) with a stochastic ODE (for infection outcomes in humans) rather than using an unacceptably slow stochastic individual level model [62, 63]. Mixed DES/ODE models can also assess the interacting effects of system capacity, service delays, screening quality, and health outcomes [64]. Inaccurate cost-benefit assessments can result from ignoring the interactions between service capacity decisions, waits and health outcomes, or by using ad hoc approximations for relevant delay times that do not fully account for disease progression and queueing in health care delivery systems.

Multiple modelling approaches can be used for model validation. For example, Chick *et al.* [52] did this as part of a study to show how long-term partnership concurrency can strongly influence the prevalence of sexually transmitted diseases. Output from differential equation, CTMC and DES models were compared to insure that the DES was coded correctly, then the DES explored how mixing patterns affected STD prevalence. Such exercises are an important part of 'structural sensitivity analysis', which can sometimes be ignored in favour of probabilistic sensitivity analysis on parameters only. It may be valuable not to define the 'best' model but rather to establish whether decisions would be different if different structures are used.

Bayesian evidence synthesis [65, 66] is being employed more commonly to inform health economic models by synthesising evidence from clinical trials with observational or other forms of study. Indirect evidence can be used to quantify parameters such as the difference in efficacy between two active treatments that have only been examined in clinical trials against placebo. The approach also allows estimation of parameters that are not directly measurable. The choice of model structure is generally driven by features of the system modelled rather than available data. These features should in turn influence the approach to evidence synthesis, in particular the parameters required by the decision model. Pragmatically of course the data available and any pre-existing evidence synthesis can also inform the parameterisation of the decision model, e.g. the choice of whether to use two separate parameters for the effectiveness of treatment, or the relative or absolute difference between those two values. Thus, this complex statistical modelling to derive the parameters does not generally impact on the choice of decision model structure as laid out in our taxonomy table but will affect the detailed model structure.

The influence of model development time on model structure selection will remain. It is often related to the experience of the modeller, the complexity of the system and the software available. If the decision problem allows, simple models with a small number of states are quicker to develop and run and easy to understand. However, where multiple risk factors and non-Markovian delays are present, an individual level model may be both quicker to code and more transparent to the decision maker. Policy makers

also need to be sensitive to the time required to provide appropriate models for the purpose required.

National decision makers such as NICE and equivalent bodies in other countries contribute heavily to methodology development, but there are also decision makers at international and more local levels, who utilise evidence on costs and benefits. NICE are interested in mean cost per QALY gained and PSA, which can appear to encourage researchers to use cohort model structures, but modellers need to be aware that PSA should not drive a decision to use an inappropriate structure. For other decision makers more detailed models may be needed to provide disaggregated output, such as cost estimates over time and variability between different geographical regions.

Finally, it is apparent that there is a great need for a wider awareness of the variety of techniques and their applicability. Many of these techniques are in the domain of Operational Research, epidemiology, and statistics. Increased training, communication and 'buying in' will help the Health Economics discipline. This is a high priority because the use of modelling and evidence synthesis for evaluating health technologies has developed substantially over the last 10 years. The use of the full range of techniques set out in the taxonomy grid and the criteria for appropriate model selection is the route by which the discipline will develop.

Acknowledgements

This paper was produced as part of a CHEBS Focus Fortnight on Patient Level Simulation Modelling hosted at the Centre for Bayesian Statistics in Health Economics (CHEBS), University of Sheffield in October 2004. The Authors would like to thank Professor Tony O'Hagan for hosting this event and encouraging the production of this paper. The Authors acknowledge each other as co-equal first authors, set out in alphabetical order.

Appendix A

The online appendix at <http://faculty.insead.edu/chick/papers/HTAModel-AppTable.html> summarises implementation issues for the models in

the taxonomy, and provides links to software reviews and vendors (the inclusion of a product or vendor is not intended to imply endorsement, and *vice versa*).

References

1. Taylor RS, Drummond MF, Salkeld G, Sullivan SD. Inclusion of cost-effectiveness in licensing requirements of new drugs: the fourth hurdle. *Br Med J* 2004; **329**: 972–975.
2. Griffin S, Claxton K, Hawkins N, Sculpher M. *Probabilistic sensitivity analysis and computationally expensive models: necessary and required?* Health Economists Study Group, Oxford, January 2005.
3. National Institute for Clinical Excellence (NICE). *Guide to the Methods of Technology Appraisal*. NICE: London, 2004.
4. Claxton K, Sculpher M, McCabe C, Briggs A, Akehurst R, Buxton M, Brazier J, O'Hagan T. Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. *Health Econ* 2005; **14**: 339–347.
5. Halpern MT, Luce BR, Brown RE, Genest B. Health and economic outcomes modeling practices: a suggested framework. *Value Health* 1998; **1**(2): 131–147.
6. Chilcott J, Brennan A, Booth A, Karnon J, Tappenden P. The role of modelling in prioritising and planning clinical trials. *Health Technol Assess* 2003; **7**(23): 1–138.
7. Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C, Luce BR. Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR task force on good research practices – modeling studies. *Value Health* 2003; **6**(1): 9–17.
8. Philips Z, Ginnelly L, Sculpher M *et al.* Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess* 2004; **8**(36): 1–172.
9. Sculpher M, Fenwick E, Claxton K. Assessing quality in decision analytic cost-effectiveness models. *Pharmacoeconomics* 2000; **17**: 461–477.
10. Karnon J. Alternative decision modelling techniques for the evaluation of health care technologies: Markov processes versus discrete event simulation. *Health Econ* 2003; **12**: 837–848.
11. Koopman JS, Jacquez G, Chick SE. Integrating discrete and continuous population modeling strategies in population health and aging: strengthening the dialog between epidemiology and demography. In *Annals of the New York Academy of Sciences*, Weinstein M, Hermalin A, Stoto MA (eds), vol. 954. New York Academy of Science: New York 2001; 268–294.

12. Barton P, Bryan S, Robinson S. Modelling in the economic evaluation of health care: selecting the appropriate approach. *J Health Serv Res Policy* 2004; **9**(2): 110–118.
13. Claxton K, Sculpher M, Drummond M. A rational framework for decision making by the National Institute for Clinical Excellence (NICE). *Lancet* 2002; **360**: 711–715.
14. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. *Med Decision Making* 1993; **13**: 322–338.
15. Briggs A, Sculpher M. An introduction to Markov modelling for economic evaluation. *Pharmacoeconomics* 1998; **13**: 397–409.
16. Eastman RC, Javitt JC, Herman WH *et al.* Model of complications of NIDDM. Model construction, assumptions. *Diabetes Care* 1997; **20**: 725–734.
17. Chilcott JB, Whitby SM, Moore R. Clinical impact and health economic consequences of posttransplant type 2 Diabetes Mellitus. *Transplant Proc* 2001; **33**(Suppl 5A): 32S–39S.
18. Brown JB, Palmer AJ, Bisgaard P, Chan W, Pedula K, Russell A. The Mt. Hood challenge: cross-testing two diabetes simulation models. *Diabetes Res Clin Prac* 2000; **50**(Suppl. 3): S57–S64.
19. Brennan A, Bansback A, Reynolds A, Conway P. Modelling the cost-effectiveness of etanercept in adults with rheumatoid arthritis in the UK. *Rheumatology* 2004; **43**: 62–72.
20. Stevenson MD, Brazier JE, Calvert NW, Lloyd-Jones M, Oakley J, Kanis JA. Description of an individual patient methodology for calculating the cost-effectiveness of treatments for osteoporosis in women. *J Oper Res Soc* 2005; **56**: 214–221.
21. Barton P, Jobanputra P, Wilson J, Bryan S, Burls A. The use of modeling to evaluate new drugs for patients with a chronic condition: the case of antibodies against tumour necrosis factor in rheumatoid arthritis. *Health Technol Assess* 2004; **8**(11): 1–104.
22. Clark W, Jobanputra P, Barton P, Burls A. The clinical and cost-effectiveness of anakinra for the treatment of rheumatoid arthritis in adults: a systematic review and economic analysis. *Health Technol Assess* 2004; **8**(18): 1–118.
23. Jacquez JA. *Compartmental Analysis in Biology and Medicine* (3rd edn). BioMedWare: Ann Arbor, MI, 1996.
24. Dangerfield BC. System dynamics applications to European health care issues. *J Oper Res Soc* 1999; **50**: 345–353.
25. Diekmann O, Heesterbeek JA. *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation*. Wiley: New York, 2000.
26. Luenberger DG. *Introduction to Dynamic Systems: Theory, Models and Applications*. Wiley: New York, 1979.
27. Kaplan EH, Craft DL, Wein LM. Emergency response to a smallpox attack: the case for mass vaccination. *Proc Nat Acad Sci* 2002; **99**(16): 10935–10940.
28. Koopman JS. Controlling smallpox. *Science* 2002; **298**: 1342–1344.
29. Kaplan EH, Wein LM. Smallpox bioterror response. *Science* 2003; **300**: 1503.
30. Bennett P, Hare A, Townsend J. Assessing the risk of vCJD transmission via surgery: models for uncertainty and complexity. *J Oper Res Soc* 2005; **56**(2): 202–213.
31. Press WH, Teukolsky SA, Vetterling WT, Flannery BP. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press: Cambridge, 1992.
32. Çinlar E. *Introduction to Stochastic Processes*. Prentice-Hall: Englewood Cliffs, NJ, 1975.
33. Banks J, Carson JS, Nelson BN. *Discrete-event System Simulation*. (2nd edn). Prentice-Hall, Inc.: Upper Saddle River, NJ, 1995.
34. Kurtz TG. Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *J Appl Probab* 1971; **8**: 344–356.
35. Pollett PK. On a model for interference between searching insect parasites. *J Austr Math Soc Ser B* 1990; **31**: 133–150.
36. Whitt W. *Stochastic-Process limits*. Springer: Berlin, 2001.
37. Brookhart MA, Hubbard AE, van der Laan ME, Colford JM, Eisenberg JNS. Statistical estimation of parameters in a disease transmission model: an analysis of a Cryptosporidium outbreak. *Stat Med* 2002; **21**(23): 3627–3638.
38. Law AM, Kelton WD. *Simulation Modeling and Analysis* (3rd edn). McGraw-Hill: New York, 2000.
39. Adams AL, Barth-Jones DC, Chick SE, Koopman JS. Simulations to evaluate HIV vaccine trial methods. *Simulation* 1998; **71**(4): 228–241.
40. Davies R, Roderick P, Raftery J. The evaluation of disease prevention and treatment using simulation models. *Eur J Oper Res* 2003; **150**(1): 53–66.
41. Fone D, Hollinghurst S, Temple M, Round A, Lester N, Weightman A, Roberts K, Coyle E, Bevan G, Palmer S. Systematic review of the use and value of computer simulation modelling in population health and health care delivery. *J Public Health Med* 2003; **25**(4): 325–335.
42. Jun JB, Jacobson SH, Swisher JR. Applications of discrete event simulation in health care clinics: a survey. *J Oper Res Soc* 1999; **50**(2): 109–123.
43. Tocher KD. *The Art of Simulation*. The English Universities Press Ltd.: London, 1963; 9–17.
44. Fetter RB, Thompson JD. The simulation of hospital systems. *Oper Res* 1965; **13**: 689–711.
45. Banks J. *Handbook of Simulation*. Wiley: New York, 1996.

46. Glynn, P. On the role of generalized semi-markov processes in simulation output analysis. In *Proceedings of the 1983 Winter Simulation Conference*, Wilson JR, Henriksen JO, Roberts SD (eds). IEEE Inc.: Piscataway, NJ, 1983; 38–42.
47. Yücesan E, Schruben LW. Structural and behavioral equivalence of simulation models. *ACM Trans Model Comput Simulation* 1992; **2**: 82–103.
48. Cooper K, Davies R, Roderick P, Chase D, Raftery J. The development of a simulation model for the treatment of coronary artery disease. *Health Care Manage Sci* 2002; **5**: 259–267.
49. Rauner MS. Using Simulation for AIDS policy modeling: benefits for HIV/AIDS prevention policy makers in Vienna, Austria. *Health Care Manage Sci* 2002; **5**(2): 121–134.
50. Davies R, O'Keefe RM, Davies HTO. Simplifying the modeling of multiple activities, multiple queueing, and interruptions, a new low level data structure. *ACM Trans Model Comp Simul* 1993; **3**(2): 332–346.
51. Halloran EM, Longini IM, Azhar N, Yang Y. Containing bioterrorist smallpox. *Science* 2002; **298**: 1428–1432.
52. Chick SE, Adams A, Koopman JS. Analysis and simulation of a stochastic, discrete-individual model of STD transmission with partnership concurrency. *Math Biosci* 2000; **166**(1): 45–68.
53. Davies R, Roderick P, Raftery J, Crabbe D, Patel P, Goddard JR. A simulation to evaluate screening for *Helicobacter pylori* infection in the prevention of peptic ulcers and gastric cancers. *Health Care Manage Sci* 2002; 249–258.
54. Oakley J. Value of information for complex cost-effectiveness models. *Research Report No. 533/02*, Department of Probability and Statistics, University of Sheffield, 2002; **5**.
55. Stevenson MD *et al.* Gaussian process modelling in conjunction with individual patient simulation modelling. A case study describing the calculation of cost-effectiveness ratios for the treatment for established osteoporosis. *Med Decision Making* 2004; **24**: 89–100.
56. O'Hagan A, Stevenson M, Madan J. Monte Carlo probabilistic sensitivity analysis for patient level simulation models. *Research Report No. 561/05*, Department of Probability and Statistics, University of Sheffield, 2005, available at <http://www.shef.ac.uk/content/1/c6/02/56/37/MCP5A.pdf>.
57. Ross SM. *Stochastic Processes* (2nd edn). Wiley: New York, 1996.
58. Babad H, Sanderson C, Naidoo B, White I, Wang D. The development of a simulation model for modeling primary prevention strategies for coronary heart disease. *Health Care Manage Sci* 2002; **2**: 69–74.
59. NHS Modernization Agency. *10 High Impact Changes for Service Improvement and Delivery: A Guide for NHS Leaders*, Leicester, 2004. [Accessed online 4 March 2005, <http://www.content.modern.nhs.uk/cmsWISE/HIC/HIC+Intro.htm>]
60. Koopman JS, Chick SE, Riolo CS, Simon CP, Jacques G. Stochastic effects on endemic infection levels of disseminating versus local contacts. *Math Biosci* 2002; **180**: 49–71.
61. Sculpher MJ, Pang FS, Manca A *et al.* Generalisability in economic evaluation studies in healthcare: a review and case studies. *Health Technol Assess* 2004; **8**(49):iii–iv, 1–192.
62. Chick SE, Soorapanth S, Koopman JS. Microbial risk assessment for drinking water. In *Operations Research and Health Care: Handbook of Methods Infection Levels of Disseminating Versus Local Contacts*. *Math Biosci* 2002; **180**: 49–71.
63. Chick SE, Soorapanth S, Koopman JS. Microbial risk assessment for drinking water. In *Operations Research and Health Care: Handbook of Methods and Applications*, Brandeau ML, Sainfort F, Pierskalla WP (eds). Kluwer Academic Publishers: Dordrecht, 2004; 467–494.
64. Güneş ED, Chick SE, Akşin OZ. Breast cancer screening: trade-offs in planning and service provision. *Health Care Manage Sci* 2004; **7**(4): 291–303.
65. Spiegelhalter DJ, Best NG. Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling. *Stat Med* 2003; **22**(23): 3687–3709.
66. Ades AE, Sculpher M, Sutton A, Abrams K *et al.* Bayesian methods for evidence synthesis in cost-effectiveness analysis. *Pharmacoeconomics* 2006; **24**(1): 1–19.